

A Verification Theorem for Indexability of Discrete Time Real State Discounted Restless Bandits

José Niño-Mora*

Department of Statistics

Carlos III University of Madrid

jnimora@alum.mit.edu, <http://alum.mit.edu/www/jnimora>

(under review, submitted November 30, 2015)

Abstract

This paper presents sufficient conditions for indexability (existence of the Whittle index) of general real state discrete time restless bandit projects under the discounted optimality criterion, with possibly unbounded reward and resource consumption functions. The main contribution is a verification theorem establishing that, if performance metrics and an explicitly defined marginal productivity (MP) index satisfy three conditions, then the project is indexable with its Whittle index being given by the MP index, in a form implying optimality of threshold policies for dynamic project control. The proof is based on partial conservation laws and infinite-dimensional linear programming duality. Further contributions include characterizations of the index as a Radon–Nikodým derivative and as a shadow price, and a recursive index-computing scheme.

Keywords: Markov decision process; discounted optimality criterion; real state; discrete time; restless bandits; optimal index policies; indexability; optimal threshold policies; infinite-dimensional linear programming; duality; conservation laws.

Contents

1	Introduction.	3
1.1	Indexability.	3
1.2	PCL-indexability approach to indexability of discrete state projects.	4
1.3	Applications of the (R)MABP with real state projects.	4
1.4	Contributions.	5
1.5	Proof roadmap for Theorem 2.1.	6
1.6	Structure of the paper.	6

*This work has been partially supported by the Spanish General Directorate for Scientific and Technical Research (DGICYT) national research projects MTM2007-63140 and MTM2010-20808. The author has presented parts of this work during its development over a number of years in talks at several conferences, including: NGI 2008, the 4th EuroNGI Conference on Next Generation Internet Networks (Krakow, Poland); NETCOOP 2009, the 3rd EuroNG Workshop on Network Control and Optimization (Eindhoven, The Netherlands); ISMP 2012, the 21st International Symposium on Mathematical Programming (Berlin, Germany); and RSME 2013, the 2013 Congress of the Spanish Royal Mathematical Society.

2	Preliminaries and formulation of the main result.	7
2.1	Notation.	7
2.2	Project control model, performance metrics and λ -price problems.	7
2.3	Indexability.	10
2.4	Strong threshold-indexability.	11
2.5	Marginal performance metrics and marginal productivity index.	11
2.6	The main result: A verification theorem for strong threshold-indexability.	13
3	Preliminary results.	14
4	LP reformulation of the λ-price problem starting from $X_0 \sim p$.	15
5	Decomposition of performance metrics.	17
6	Partial conservation laws and a relaxation of $\tilde{L}_\lambda(x)$.	20
7	Performance metrics as functions of the threshold.	22
7.1	Càdlàg property.	22
7.2	Monotonicity of resource metric.	24
7.3	Analysis of discontinuities.	24
7.4	Bounded variation and MP index as Radon–Nikodým derivative.	26
7.5	Relations between marginal performance metrics.	29
8	Dual LP and strong duality.	30
9	Optimality of threshold policies.	35
10	Further relations between MP metrics.	36
11	Proof of Theorem 2.1.	37
12	Examples.	37
12.1	Optimal stopping and Gittins index.	37
12.2	Optimal dynamic transmission over a noisy channel.	38
13	Concluding remarks.	40
A	Geometric and economic interpretations of the MP index.	40
B	Checking condition (PCLI3).	44
B.1	Piecewise differentiable $F(x, \cdot)$ and $G(x, \cdot)$	44
B.2	Piecewise constant $G(x, \cdot)$	46
C	Performance metrics and index computation.	47

1 Introduction.

This paper considers a class of Markov decision processes (MDPs) known as *bandits*, which are models for the optimal dynamic allocation of a resource to a stochastic *project* admitting two modes of operation, called *active* and *passive*. We will hence use the terms “bandit” and “project” interchangeably. The focus will be on the discrete time infinite horizon real state case, under the discounted optimality criterion.

We thus consider a project whose state X_t evolves across the state space X , which is an interval of the real line \mathbb{R} . At the start of period $t = 0, 1, \dots$, the state X_t is observed and then a *control action* $A_t \in \{0, 1\}$ is selected, where $A_t = 1$ (resp. $A_t = 0$) models the project being operated in the active (resp. passive) mode. If $(X_t, A_t) = (x, a)$ then (i) the project is allocated $c(x, a)$ resource units and yields a reward $r(x, a)$ in period t , both of which being discounted with factor $\beta \in [0, 1)$; and (ii) the state moves to X_{t+1} according to a Markov transition law $\kappa^a(\cdot, \cdot)$. Both the reward and the resource consumption functions $r(x, a)$ and $c(x, a)$ may be unbounded, under certain assumptions.

Actions are selected through a *control policy* $\pi = (\pi_t)_{t=0}^\infty$ from the class Π of *admissible policies*, which are history-dependent randomized. Suppose that a price λ is charged per unit resource used, so $V_\lambda(x, \pi) \triangleq \mathbb{E}_x^\pi[\sum_{t=0}^\infty \{r(X_t, A_t) - \lambda c(X_t, A_t)\} \beta^t]$ is the expected net project value starting from state x under policy π . Consider the following parametric collection of λ -price problems: for $\lambda \in \mathbb{R}$,

$$P_\lambda: \quad \text{find } \pi_\lambda^* \in \Pi \text{ such that } V_\lambda(x, \pi_\lambda^*) = V_\lambda^*(x) \triangleq \sup_{\pi \in \Pi} V_\lambda(x, \pi) \text{ for each state } x. \quad (1)$$

Note that that a policy π_λ^* is P_λ -optimal if $V_\lambda(x, \pi_\lambda^*) = V_\lambda^*(x)$ for each state x .

1.1 Indexability.

The project (or, more properly, the problem collection $\{P_\lambda: \lambda \in \mathbb{R}\}$) is called *indexable* with *index* $\lambda^*: \mathsf{X} \rightarrow \mathbb{R}$ if, for every price λ and state x : (i) the active action is P_λ -optimal in state x if and only if $\lambda^*(x) \geq \lambda$; and (ii) the passive action is P_λ -optimal in state x if and only if $\lambda^*(x) \leq \lambda$.

Such an *indexability* property has played a fundamental role in the study of bandit models. It was first identified in *classic bandits*, in which the state does not change under the passive action, with the resource consumption function being $c(x, a) \equiv a$, in the pioneering work of Bradt et al. [5] and Bellman [2] on models for the sequential design of experiments. Classic bandits are generally indexable. In a celebrated result, Gittins and Jones [14] used such an index to obtain an optimal policy for the long-standing *multi-armed bandit problem* (MABP), involving a finite collection of projects, one of which must be active at each time, under the infinite horizon discounted reward criterion. They showed that the *priority index rule* that engages at each time a project of currently largest index value is optimal. See also Gittins [12, 13]. The index of classic bandits, known as the *Gittins index*, and its multiple applications, have been the subject of extensive research attention. See, e.g., the surveys Bergemann and Välimäki [3], Mahajan and Teneketzis [33], and Weber [52].

The concept of indexability was extended in the seminal work of Whittle [56] from classic to *restless bandits*, which can change state while passive, in which case the literature properly refers to the *Whittle index*. Whittle [56] considers an extension of the MABP where projects are restless and a fixed number of them must be active at each time, a version of the generally intractable (see Papadimitriou and Tsitsiklis [43]) *restless MABP* (RMABP), motivating it with several applications such as aircraft surveillance and dynamic worker selection. If the projects are all indexable, Whittle proposed to give higher priority for being engaged to those with currently larger index values. Although such a *Whittle index policy* is generally suboptimal for the RMABP, it possesses (see Weber and Weiss [53]) a form

of asymptotic optimality under certain conditions. Yet, as pointed out in Whittle [56], unlike classic bandits, restless bandits need not be indexable. Further, Whittle [56] considers a relaxed problem whose optimal policy is given in terms of individual project indices under indexability, and which can be used to bound the original problem’s optimal value.

Due to the broad modeling power of the RMABP and the practical appeal of Whittle’s index policy, a large body of work has investigated over the last quarter century its application to a wide variety of models, often reporting a nearly optimal performance in simulation studies. For a sample of such work see, e.g., Whittle [57, Ch. 14.7], Veatch and Wein [49], Raissi-Dehkordi and Baras [44], Dusonchet and Hongler [11], Caro and Gallien [6], Caro and Yoo [7], Huberman and Wu [21], Kumar and Saranga [24], Temple and Frazzoli [47], Wei et al. [54], and Wang and Chen [50].

To apply the Whittle index policy to a particular RMABP model, two roadblocks must be overcome: the first is to establish indexability for the constituent projects; the second is to devise an efficient index-computing scheme. Typically, researchers have tackled both issues through ad hoc approaches that exploit the special structure of the model at hand.

1.2 PCL-indexability approach to indexability of discrete state projects.

In contrast, a general approach to resolve such roadblocks in discrete state projects has been developed in Niño-Mora [34, 35, 36, 37], giving sufficient indexability conditions for a generic project, which can be checked analytically or numerically, along with an index algorithm. Their validity follows from satisfaction by project performance metrics of *partial conservation laws* (PCLs), also introduced in that work, whence we refer to *PCL-indexability conditions*. PCLs are an extension of the conservation laws satisfied by certain stochastic scheduling models, including the MABP, which have been used to establish the optimality of index policies. See, e.g., Coffman and Mitrani [9], Shanthikumar and Yao [46], Bertsimas and Niño-Mora [4], and the survey Niño-Mora [40].

1.3 Applications of the (R)MABP with real state projects.

While most work to date on applications of the (R)MABP has considered discrete state models, the study of models with real state projects has attracted increasing attention over the last decade. Washburn [51] surveys applications of the (R)MABP to Partially Observable MDP (POMDP) sensor management models, noting the difficulties caused by real state spaces and stating that “part of the difficulty of applying restless bandit methods may lie in the difficulty of checking Whittle’s indexability conditions ... and the likelihood that these conditions are not satisfied for typical sensor management problems.”

Even in the classic case, computation of the Gittins index raises significant challenges for real state projects. Krishnamurthy and Evans [22] consider a POMDP multi-target tracking model formulated as a MABP with real state projects, and give a value iteration scheme to approximately compute the index. Krishnamurthy and Wahlberg [23] extend such results to a general POMDP setting, and give more efficient algorithms for the Gittins index under certain assumptions.

La Scala and Moran [25] formulate the problem of dynamic beam scheduling of a phased array radar to minimize the tracking error of two targets as a RMABP with real state projects with deterministic dynamics. They point out the lack of methods to address the indexability issue in such a setting, and focus on identifying conditions under which a one-step greedy policy is optimal.

Le Ny et al. [26] and Liu and Zhao [29, 30] address equivalent RMABP models, though motivated by different applications. Le Ny et al. [26] consider the optimal dynamic allocation of m mobile sensors to track $n > m$ targets with independent binary state Markov dynamics. A target’s state

is known only when sensed. Otherwise, only the *information state* is known, giving the probability that the target is in one of the states. A POMDP model is considered, which is reformulated as a RMABP where the projects (targets) have as state space the interval $[0, 1]$ on which the information state lies. Indexability is established under the discounted criterion by exploiting special properties of the optimal value function for the single-project λ -price subproblems (1), and the Whittle index is evaluated in closed form. Liu and Zhao [29, 30] consider the optimal dynamic allocation of m sensors to sense and transmit packets through n parallel communication channels evolving as binary state (good or bad) Markov chains, with the channel state being known only when sensed. While the model is equivalent to that in Le Ny et al. [26], Liu and Zhao [30] establish indexability, under the discounted and the long-run average criteria, through a different approach based on bounding a project's total passive time under an optimal policy. Both Le Ny et al. [26] and Liu and Zhao [29, 30] report a nearly optimal performance for the Whittle index policy.

Liu et al. [32] establish indexability for a class of restless bandits where the active action resets the system evolution, extending results in Le Ny et al. [26] and in Liu and Zhao [30].

Le Ny et al. [27] consider a continuous time RMABP model subject to Kalman filter dynamics, and establish indexability in the case of real state projects by exploiting the model's structure. Simulation results are reported showing that the Whittle index policy is near optimal.

1.4 Contributions.

The conference papers Niño-Mora [38, 39] introduce an extension to real state projects of the general PCL-based sufficient indexability conditions and index evaluation scheme developed in Niño-Mora [34, 35, 36, 37] for discrete state projects, demonstrating through numerical experiments their applicability to the models considered in Liu and Zhao [30] and Liu et al. [31]. See also Niño-Mora and Villar [41, 42]. Yet, no proofs establishing the validity of the proposed indexability conditions for real state projects are given in such work.

Motivated by the above, the major goal of this paper is to provide a general approach for establishing indexability and evaluating the index of a project, by laying down a rigorous foundation to the general PCL-based sufficient indexability conditions and index evaluation scheme for real state projects outlined in the above referenced work, with proofs for all results.

Our main result is the *verification theorem* in Theorem 2.1, ensuring the indexability of a project under certain conditions with its Whittle index being given by a so-called *marginal productivity* (MP) index, which, unlike the Whittle index, is explicitly defined. Such a result establishes indexability in a form that implies the optimality of *threshold policies* for the λ -price subproblems (1), motivated by the simplicity and prevalence of such policies (cf. Heyman and Sobel [20, Ch. 8]).

In short, we consider certain *reward* and *resource (usage) performance metrics* for a project, denoted respectively by $F(x, z)$ and $G(x, z)$, where x is the initial project state and z is the threshold, representing the policy that takes the active action in states larger than z and the passive one otherwise. We further consider corresponding *marginal reward* and *marginal resource* metrics, denoted respectively by $f(x, z)$ and $g(x, z)$, as well as the *marginal MP* metric $m(x, z) \triangleq f(x, z)/g(x, z)$, which is defined when $g(x, z) \neq 0$. The *MP index* is given by $m^*(x) \triangleq m(x, x)$.

The *PCL-indexability conditions* we consider are: (PCLI1) the marginal resource metric $g(x, z)$ is positive for every state x and threshold z ; (PCLI2) the MP index m^* is monotone nondecreasing, continuous and bounded below; and (PCLI3) the reward and resource metrics are related by

$$F(x, z_2) - F(x, z_1) = \int_{(z_1, z_2]} m^*(z) G(x, dz), \quad z_1 < z_2,$$

so, for each state x , and viewed as functions of the threshold variable z , $F(x, z)$ is an indefinite *Lebesgue–Stieltjes* (LS) *integral* (see Carter and van Brunt [8]) of m^* with respect to $G(x, z)$.

Theorem 2.1 states that a project satisfying (PCLI1–PCLI3) is indexable with index m^* . This result reduces the task of establishing a project’s indexability to the *performance analysis* problem of verifying certain properties of project metrics under threshold policies, and further provides an explicit expression for the index.

Other contributions of the paper include the following. Proposition 7.1 characterizes the MP index as a *Radon–Nikodým derivative* under PCL-indexability.

The geometric and economic interpretation of the MP index is clarified in Proposition A.1, which characterizes the MP index as a *resource shadow price* under PCL-indexability. This extends corresponding results in Niño-Mora [35, 36] for discrete state projects.

Note that the PCL-indexability conditions developed in Niño-Mora [34, 35, 36, 37] for discrete state projects correspond to (PCLI1, PCLI2) only, which are shown there to imply (PCLI3). In the present setting, since (PCLI3) may be hard to verify, we give Propositions B.1 and B.2, which ensure that (PCLI3) holds provided (PCLI1, PCLI2) and some additional simpler conditions hold.

Further, Proposition C.1 gives a recursion for computing the MP index in cases where it cannot be evaluated in closed form.

1.5 Proof roadmap for Theorem 2.1.

The proof of Theorem 2.1 is complex, drawing on and integrating diverse techniques, with the bulk of the paper being devoted to groundwork. For clarity, the proof is presented as following from a series of lemmas. We next outline a high level roadmap, highlighting the key ideas.

A major tool in the proof is the use of *infinite-dimensional linear programming* (LP). See Anderson and Nash [1]. The starting point is to reformulate a version $P_\lambda(x)$ of the λ -price problem P_λ in (1) with a fixed initial state x as an infinite-dimensional LP problem $L_\lambda(x)$.

We then use $L_\lambda(x)$ to derive a collection of *decomposition identities for performance metrics*. These are used to formulate an equivalent problem $\tilde{P}_\lambda(x)$ to $P_\lambda(x)$ with zero active rewards, which is convenient for the analyses, and whose LP reformulation is denoted by $\tilde{L}_\lambda(x)$.

The decomposition identities for the resource metric $G(x, \pi)$ are used to formulate a collection of PCLs on certain project metrics, which are shown to be satisfied under (PCLI1). Such PCLs directly yield an *LP relaxation* $R_\lambda(x)$ of problem $\tilde{L}_\lambda(x)$, which plays a central role in the proof.

We characterize the optimal solutions to $R_\lambda(x)$ under PCL-indexability, showing that they are achieved by threshold policies, which yields their optimality for the λ -price problem. This result is obtained via *LP duality*, by postulating, justifying and solving in closed form a dual LP to $R_\lambda(x)$, and establishing *strong duality*. The MP index m^* is the key ingredient of the optimal dual solution.

A characterization of optimal threshold policies for the λ -price problem is obtained from such results, which is shown to imply indexability with index m^* .

1.6 Structure of the paper.

The remainder of the paper is organized as follows. §2 sets the notation used throughout, and presents required preliminary material to properly formulate the main result in §2.6. §3 presents preliminary results shedding light on the PCL-indexability conditions. §4 gives the LP reformulation $L_\lambda(x)$ of the problem $P_\lambda(x)$ referred to above. §5 presents the decomposition of the project’s performance metrics outlined above. §6 shows that the project metrics satisfy certain PCLs, and uses them to obtain an LP relaxation $R_\lambda(x)$ of problem $\tilde{L}_\lambda(x)$. §7 establishes certain properties of performance

metrics as functions of the threshold variable, which are used in subsequent analyses, and further characterizes the MP index of a PCL-indexable project as a Radon–Nikodým derivative. §8 sets out to solve the LP relaxation $R_\lambda(x)$ under PCL-indexability, via LP duality, establishing strong duality. §9 establishes, under PCL-indexability, the optimality of threshold policies for each λ -price problem and further characterizes in terms of the MP index m^* the optimal threshold policies. §10 derives further relations between MP metrics resulting from the above work, which play a key role in the indexability analysis. §11 draws on the above to give the proof of Theorem 2.1. §12 outlines the application of the PCL-indexability conditions to some examples. Finally, §13 concludes.

Three appendices provide relevant ancillary material. §A presents geometric and economic interpretations of the MP index under PCL-indexability, showing in particular that it can be characterized as a *resource shadow price*. §B gives practical tools to establish that condition (PCL13) holds. §C gives recursions for the numerical approximation of the project’s index and performance metrics in cases where they cannot be evaluated in closed form.

2 Preliminaries and formulation of the main result.

This section describes the single-project optimal control model that is our prime focus, and presents key concepts required to formulate our main result, Theorem 2.1.

2.1 Notation.

We start by setting the notation used in the paper. Given endpoints $z_1 \leq z_2$ in the extended real numbers $\overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{-\infty, \infty\}$, we denote by (z_1, z_2) , $[z_1, z_2)$, $(z_1, z_2]$ and $[z_1, z_2]$ the corresponding real intervals. Thus, e.g., $[z_1, z_2] \triangleq \{z \in \mathbb{R} : z_1 \leq z \leq z_2\}$, so $[-\infty, \infty] = (-\infty, \infty) = \mathbb{R}$.

We denote by $X \triangleq [\ell, u]$ a closed real interval with possibly infinite endpoints $-\infty \leq \ell < u \leq \infty$, and by $\mathcal{B}(X)$ its Borel σ -algebra. For $S \subseteq X$, we write $S^c \triangleq X \setminus S$, and denote by 1_S its indicator function. We sometimes write $1_{(z, u]}(x)$ as $1_{\{x > z\}}$, and similarly with other intervals.

For a two-variable function $h : X \times Y \rightarrow \mathbb{R}$, where Y is a subset of \mathbb{R} , we write as $h(\cdot, y) : X \rightarrow \mathbb{R}$ and $h(x, \cdot) : Y \rightarrow \mathbb{R}$ the single-variable functions obtained by fixing the value of one variable, and write $\Delta_{x=x_0}^{x=x_1} h(x, y) \triangleq h(x_1, y) - h(x_0, y)$ and $\Delta_{y=y_0}^{y=y_1} h(x, y) \triangleq h(x, y_1) - h(x, y_0)$. We denote by $|h|$ the absolute value of h , so $|h|(x, y) \triangleq |h(x, y)|$. If $h(\cdot, y)$ has finite left- and right limits at x_0 , denoted by $h(x_0^-, y)$ and $h(x_0^+, y)$, we write as $\Delta_1 h(x_0, y) \triangleq h(x_0^+, y) - h(x_0^-, y)$ the *jump* of $h(\cdot, y)$ at x_0 . The notation $h(x, y_0^-)$, $h(x, y_0^+)$ and $\Delta_2 h(x, y_0) \triangleq h(x, y_0^+) - h(x, y_0^-)$ is interpreted similarly.

We denote by $\mathbb{C}(X)$ the space of real-valued continuous functions on X . We will further refer to the space of real-valued *càdlàg* (“continue à droite, avec des limites à gauche,” i.e., right-continuous with left limits) functions on X .

We denote by $\int_B J d\alpha$ or $\int_B J(x) \alpha(dx)$ the LS integral (see Carter and van Brunt [8]) of a function J with respect to a function α over a set B . When B is the domain of J we write $\int J d\alpha$.

Throughout the paper, the terms “increasing” and “decreasing” are used in the strict sense, and “iff” stands for “if and only if.”

2.2 Project control model, performance metrics and λ -price problems.

We consider a general restless bandit model for the optimal dynamic allocation of a resource to a project, whose state X_t evolves in discrete time over an infinite horizon across the state space $X \triangleq [\ell, u]$ (see §2.1).

At the start of each time period $t = 0, 1, \dots$, the project controller observes the state X_t and then selects a *control action* $A_t \in \{0, 1\}$, where $A_t = 1$ (resp. $A_t = 0$) models the project being operated in the *active mode* (resp. *passive mode*) in period t . We denote by $\mathbf{K} \triangleq \mathbf{X} \times \{0, 1\}$ the state-action space. If $(X_t, A_t) = (x, a)$ then: (i) the project is allocated $c(x, a)$ resource units and yields a reward $r(x, a)$ in period t , both of which being geometrically discounted with factor $\beta \in [0, 1]$; and (ii) the state moves to X_{t+1} according to a Markov transition kernel $\kappa^a(\cdot, \cdot)$ on \mathbf{X} with cumulative distribution function (CDF) $Q^a(x, \cdot)$, so $\mathbf{P}\{X_{t+1} \leq y \mid (X_t, A_t) = (x, a)\} = Q^a(x, y) \triangleq \kappa^a(x, (-\infty, y])$.

The following assumption is maintained throughout the paper.

Assumption 2.1. *For each action a :*

(i) $r(\cdot, a), c(\cdot, a) \in \mathbb{C}(\mathbf{X})$, with $c(\cdot, a)$ satisfying

$$0 \leq c(x, 0) < c(x, 1), \quad x \in \mathbf{X}; \quad (2)$$

(ii) *there exists a measurable weight function $w: \mathbf{X} \rightarrow [1, \infty)$ and constants $M > 0$, $\gamma \in [\beta, 1)$ such that, for every state x ,*

$$(ii.a) \quad \max\{|r|(x, a), c(x, a)\} \leq M w(x);$$

$$(ii.b) \quad \beta \tilde{w}(x, a) \leq \gamma w(x), \text{ where } \tilde{w}(x, a) \triangleq \int w(y) Q^a(x, dy);$$

Note that Assumption 2.1(ii) corresponds to the *weighted supremum norm* approach to discounted MDPs (see Heilmann [15]; Lippman [28]; Van Nunen and Wessels [48]; Wessels [55] and Hernández-Lerma and Lasserre [19, Ass. 8.3.2]) which allows $r(\cdot, a)$ and $c(\cdot, a)$ to be unbounded.

Actions are selected through a *control policy* $\pi = (\pi_t)_{t=0}^\infty$ from the class Π of *history-dependent randomized policies*, which at each time t selects the action $A_t = a$ with probability $\pi_t(a \mid \mathcal{H}_t)$, a measurable function of the history $\mathcal{H}_t \triangleq ((X_s, A_s)_{s=0}^{t-1}; X_t)$. We further consider the class Π^{SR} of *stationary randomized policies*, where $\pi_t(a \mid \mathcal{H}_t) = p(a \mid X_t)$ for some function $p(\cdot \mid \cdot) \geq 0$ with $\sum_a p(a \mid x) = 1$ for each x , and the class Π^{SD} of *stationary deterministic policies*, where $p(1 \mid x) = 1_B(x)$ for some *active (state) region* $B \in \mathcal{B}(\mathbf{X})$. Thus, we will refer to the *B-policy*, writing $\pi = B$.

For an *admissible policy* $\pi \in \Pi$ and a probability distribution p on $\mathcal{B}(\mathbf{X})$ for drawing the initial state ($X_0 \sim p$), let \mathbf{P}_p^π denote the probability distribution they and the transition law determine (by the Ionescu Tulcea extension theorem; see Hernández-Lerma and Lasserre [18, p. 178, Prop. C.10]) on the space \mathbf{K}^∞ of state-action paths with the product σ -algebra. We denote by \mathbf{E}_p^π the corresponding expectation, writing \mathbf{P}_x^π and \mathbf{E}_x^π when $p = \delta_x$, the Dirac measure concentrated at x .

We evaluate a policy π starting from p by the *reward* and *resource (usage) performance metrics*

$$F(p, \pi) \triangleq \mathbf{E}_p^\pi \left[\sum_{t=0}^\infty \beta^t r(X_t, A_t) \right] \quad \text{and} \quad G(p, \pi) \triangleq \mathbf{E}_p^\pi \left[\sum_{t=0}^\infty \beta^t c(X_t, A_t) \right], \quad (3)$$

which we write as $F(x, \pi)$ and $G(x, \pi)$ when $X_0 = x$, and as $F(p, B)$ and $G(p, B)$ when $\pi = B$.

Suppose that a price λ is charged per unit resource used, and denote by $V_\lambda(p, \pi) \triangleq F(p, \pi) - \lambda G(p, \pi)$ the resulting *net project value metric*. Consider the following collection of optimal project control problems, which is parameterized by the resource price $\lambda \in \mathbb{R}$:

$$P_\lambda: \quad \text{find } \pi_\lambda^* \in \Pi \text{ such that } V_\lambda(x, \pi_\lambda^*) = \sup_{\pi \in \Pi} V_\lambda(x, \pi) \text{ for each } x \in \mathbf{X}. \quad (4)$$

We call P_λ the project's λ -price problem and denote by $V_\lambda^*(x) \triangleq \sup_{\pi \in \Pi} V_\lambda(x, \pi)$ its *optimal value function*, for $x \in \mathbf{X}$. We say that a policy π_λ^* is P_λ -*optimal* if $V_\lambda(x, \pi_\lambda^*) = V_\lambda^*(x)$ for each x .

We next outline relevant results from the weighted supremum norm approach to MDPs, which in particular ensure that each λ -price problem can be solved by a stationary deterministic policy. Let $\mathbb{B}_w(\mathbf{K})$ and $\mathbb{B}_w(\mathbf{X})$ denote the Banach spaces (see Hernández-Lerma and Lasserre [19, Prop. 7.2.1]) of measurable w -bounded functions $u: \mathbf{K} \rightarrow \mathbb{R}$ and $v: \mathbf{X} \rightarrow \mathbb{R}$, respectively, having finite w -norms

$$\|u\|_w \triangleq \sup \left\{ \frac{|u|(x, a)}{w(x)} : (x, a) \in \mathbf{K} \right\} \quad \text{and} \quad \|v\|_w \triangleq \sup \left\{ \frac{|v|(x)}{w(x)} : x \in \mathbf{X} \right\}, \quad (5)$$

and write as $\mathbb{P}_w(\mathbf{X})$ the space of probability measures p on $\mathcal{B}(\mathbf{X})$ with finite w -norm

$$\|p\|_w \triangleq \int w \, dp = \mathbb{E}_p[w(X_0)]. \quad (6)$$

We further denote by $\mathbb{C}_w(\mathbf{K}) \triangleq \mathbb{C}(\mathbf{K}) \cap \mathbb{B}_w(\mathbf{K})$ and $\mathbb{C}_w(\mathbf{X}) \triangleq \mathbb{C}(\mathbf{X}) \cap \mathbb{B}_w(\mathbf{X})$ the corresponding subspaces of w -bounded continuous functions.

Henceforth, $F(\cdot, \pi)$, $G(\cdot, \pi)$ and $V_\lambda(\cdot, \pi)$ denote the functions on \mathbf{X} corresponding to the reward, resource and net value metrics $F(x, \pi)$, $G(x, \pi)$ and $V_\lambda(x, \pi)$.

Remark 2.1. Under Assumption 2.1 (cf. Hernández-Lerma and Lasserre [19, The. 8.3.6]):

- (i) $r(\cdot, a), c(\cdot, a) \in \mathbb{C}_w(\mathbf{X})$; $r, c \in \mathbb{C}_w(\mathbf{K})$; and $\kappa^a(x, \cdot) \in \mathbb{P}_w(\mathbf{X})$.
- (ii) For any $\pi \in \Pi$ and $\lambda \in \mathbb{R}$, $F(\cdot, \pi), G(\cdot, \pi), V_\lambda(\cdot, \pi) \in \mathbb{B}_w(\mathbf{X})$; letting $M_\gamma \triangleq M/(1 - \gamma)$, we have

$$\max\{\|F(\cdot, \pi)\|_w, \|G(\cdot, \pi)\|_w\} \leq M_\gamma, \quad \|V_\lambda(\cdot, \pi)\|_w \leq (1 + |\lambda|)M_\gamma. \quad (7)$$

- (iii) For any $\pi \in \Pi$, $p \in \mathbb{P}_w(\mathbf{X})$ and $\lambda \in \mathbb{R}$, $F(p, \pi)$, $G(p, \pi)$ and $V_\lambda(p, \pi)$ are well defined and finite, being equal to $\int F(x, \pi) p(dx)$, $\int G(x, \pi) p(dx)$ and $\int V_\lambda(x, \pi) p(dx)$, respectively, and satisfy

$$\max\{|F|(p, \pi), |G|(p, \pi)\} \leq M_\gamma \|p\|_w, \quad |V_\lambda|(p, \pi) \leq (1 + |\lambda|)M_\gamma \|p\|_w. \quad (8)$$

- (iv) For an active region $B \in \mathcal{B}(\mathbf{X})$, $F(\cdot, B)$ and $G(\cdot, B)$ are characterized (cf. Hernández-Lerma and Lasserre [19, Remark 8.3.10]) as the unique solutions in $\mathbb{B}_w(\mathbf{X})$ to the fixed-point equations

$$\begin{aligned} F(x, B) &= r(x, 1_B(x)) + \beta \int F(y, B) Q^{1_B(x)}(x, dy), \quad x \in \mathbf{X} \\ G(x, B) &= c(x, 1_B(x)) + \beta \int G(y, B) Q^{1_B(x)}(x, dy), \quad x \in \mathbf{X}. \end{aligned} \quad (9)$$

Later (see §5) we will find it convenient to use the reformulation of (9) in terms of the bounded linear operator $\mathcal{L}^*: \mathbb{B}_w(\mathbf{X}) \rightarrow \mathbb{B}_w(\mathbf{K})$ mapping v to

$$\mathcal{L}^*v(x, a) \triangleq v(x) - \beta \int v(y) Q^a(x, dy), \quad (x, a) \in \mathbf{K}, \quad (10)$$

namely,

$$\begin{aligned} \mathcal{L}^*F(\cdot, B)(x, 1_B(x)) &= r(x, 1_B(x)), \quad x \in \mathbf{X} \\ \mathcal{L}^*G(\cdot, B)(x, 1_B(x)) &= c(x, 1_B(x)), \quad x \in \mathbf{X}. \end{aligned} \quad (11)$$

(v) For any price $\lambda \in \mathbb{R}$, the *Bellman operator* $\mathcal{T}_\lambda: \mathbb{B}_w(\mathbf{X}) \rightarrow \mathbb{B}_w(\mathbf{X})$ defined by

$$\mathcal{T}_\lambda v(x) \triangleq \max_{a \in \{0,1\}} \mathcal{T}_\lambda^a v(x), \quad \text{where } \mathcal{T}_\lambda^a v(x) \triangleq r(x, a) - \lambda c(x, a) + \beta \int v(y) Q^a(x, dy), \quad (12)$$

is a *contraction mapping* with modulus γ (see Assumption 2.1(ii)) and the *Bellman equation* (BE)

$$v = \mathcal{T}_\lambda v, \quad (13)$$

has a unique fixed point in $\mathbb{B}_w(\mathbf{X})$ given by the optimal value function V_λ^* of problem P_λ , with

$$\|V_\lambda^*\|_w \leq (1 + |\lambda|)M_\gamma. \quad (14)$$

(vi) For any $\lambda \in \mathbb{R}$ there exists a P_λ -optimal stationary deterministic policy, with a policy in Π^{SD} being P_λ -optimal iff it selects in each state x an action a with $\mathcal{T}_\lambda^a V_\lambda^*(x) \geq \mathcal{T}_\lambda^{1-a} V_\lambda^*(x)$.

In light of Remark 2.1(vi), we will say that *action a is P_λ -optimal in state x* if $\mathcal{T}_\lambda^a V_\lambda^*(x) \geq \mathcal{T}_\lambda^{1-a} V_\lambda^*(x)$. Hence, both actions are P_λ -optimal in x iff the following break-even equation holds:

$$\mathcal{T}_\lambda^0 V_\lambda^*(x) = \mathcal{T}_\lambda^1 V_\lambda^*(x). \quad (15)$$

2.3 Indexability.

We will address the problem collection $\{P_\lambda: \lambda \in \mathbb{R}\}$ through the concept of *indexability*, extended by Whittle [56] from its origin in classic bandits to restless bandits with resource consumption $c(x, a) \equiv a$, and further extended in Niño-Mora [35] to general $c(x, a)$.

For a price λ and action a , define the P_λ -optimal a -set $S_\lambda^{*,a}$ by

$$S_\lambda^{*,a} \triangleq \{x \in \mathbf{X}: \mathcal{T}_\lambda^a V_\lambda^*(x) \geq \mathcal{T}_\lambda^{1-a} V_\lambda^*(x)\}, \quad (16)$$

so it consists of those states x where action a is P_λ -optimal. Under indexability, such sets are characterized by an *index* attached to project states.

Definition 2.1 (Indexability). We call the project *indexable* if there exists an *index* $\lambda^*: \mathbf{X} \rightarrow \mathbb{R}$ such that

$$S_\lambda^{*,1} = \{x \in \mathbf{X}: \lambda^*(x) \geq \lambda\} \quad \text{and} \quad S_\lambda^{*,0} = \{x \in \mathbf{X}: \lambda^*(x) \leq \lambda\}, \quad \lambda \in \mathbb{R}. \quad (17)$$

We refer to the *Gittins index* and to the *Whittle index* depending on whether the project is classic (no passive transitions) or restless.

Remark 2.2. (i) The project is indexable with index λ^* if the P_λ -optimal action in each state is monotone nonincreasing in λ , with both actions being P_λ -optimal in x only for $\lambda = \lambda^*(x)$; thus, for $\lambda < \lambda^*(x)$ (resp. $\lambda > \lambda^*(x)$), $a = 1$ (resp. $a = 0$) is the only P_λ -optimal action in x .

(ii) Indexability can be defined as the following property of the function $J_\lambda^* \triangleq \mathcal{T}_\lambda^1 V_\lambda^* - \mathcal{T}_\lambda^0 V_\lambda^*$: the project is indexable with index λ^* if, for each state x , the equation $J_\lambda^*(x) = 0$ in the variable λ has the unique root $\lambda = \lambda^*(x)$, with $J_\lambda^*(x) > 0$ for $\lambda < \lambda^*(x)$ and $J_\lambda^*(x) < 0$ for $\lambda > \lambda^*(x)$.

As noted in Whittle [56], not all restless projects are indexable. This raises the issue of finding sufficient conditions for indexability that do not require evaluation of the optimal value function V_λ^* in terms of which indexability is defined. That is the prime goal of this paper.

2.4 Strong threshold-indexability.

We will focus the analysis of indexability on the case where it holds consistently with *threshold policies*, motivated by their simplicity and prevalence in applications. To each threshold $z \in \overline{\mathbb{R}}$ we associate two stationary deterministic policies: the z -policy, with active region $(z, u]$ (see §2.1), so the project is active in period t ($A_t = 1$) iff its state is above threshold ($X_t > z$); and the z^- -policy, with active region $[z, u]$, so $A_t = 1$ iff $X_t \geq z$. We write the corresponding metrics as $F(p, z)$, $G(p, z)$, $F(p, z^-)$ and $G(p, z^-)$.

We denote by $F(p, \cdot)$ and $G(p, \cdot)$ the functions on \mathbb{R} mapping z to $F(p, z)$ and $G(p, z)$, respectively. We will resolve the apparent ambiguity of the notation $F(p, z^-)$ and $G(p, z^-)$ in Lemma 7.2, which ensures that $F(p, z^-)$ and $G(p, z^-)$ are also the left limits at z of $F(p, \cdot)$ and $G(p, \cdot)$.

Definition 2.2 (Strong threshold-indexability). We call a project *strongly threshold-indexable* if it is indexable and, for every price $\lambda \in \mathbb{R}$, there exists a threshold $z \in \overline{\mathbb{R}}$ such that both the z -policy and the z^- -policy are P_λ -optimal.

Further, for a threshold $z \in \overline{\mathbb{R}}$ and a scalar $\alpha \in [0, 1]$ we will refer to the *randomized threshold policy* $\alpha z^- + (1 - \alpha)z$, which, in each state x , takes action $a = 1$ if $x > z$; takes action $a = 0$ if $x < z$; and takes actions $a = 0$ and $a = 1$ with probabilities α and $1 - \alpha$, respectively, if $x = z$.

Remark 2.3.

- (i) It is easily verified that the performance metrics for the policy $\alpha z^- + (1 - \alpha)z$ satisfy

$$\begin{aligned} F(p, \alpha z^- + (1 - \alpha)z) &= \alpha F(p, z) + (1 - \alpha)F(p, z^-) \\ G(p, \alpha z^- + (1 - \alpha)z) &= \alpha G(p, z) + (1 - \alpha)G(p, z^-). \end{aligned} \tag{18}$$

- (ii) If for the model of concern the relevant threshold policies take the active action in low states, one can apply the present framework to the model with state $\hat{X}_t \triangleq -X_t$.

2.5 Marginal performance metrics and marginal productivity index.

We will use the concepts of *marginal performance metrics* and *marginal productivity* (MP) index, introduced in Niño-Mora [35] for discrete state bandits, which are adapted next to the present real state setting.

For an action a and a policy π , let $\langle a, \pi \rangle$ denote the policy that takes action a in period $t = 0$ and then switches to π from $t = 1$ onwards. We measure the changes in reward earned and in resource expended, respectively, resulting from a change from policy $\langle 0, \pi \rangle$ to $\langle 1, \pi \rangle$ starting from state x by the *marginal reward metric* (see §2.1 for the notation $\Delta_{a=0}^{a=1}$)

$$f(x, \pi) \triangleq \Delta_{a=0}^{a=1} F(x, \langle a, \pi \rangle) = F(x, \langle 1, \pi \rangle) - F(x, \langle 0, \pi \rangle) \tag{19}$$

and the *marginal resource (usage) metric*

$$g(x, \pi) \triangleq \Delta_{a=0}^{a=1} G(x, \langle a, \pi \rangle). \tag{20}$$

Remark 2.4. (i) It follows from Remark 2.1(ii) and (19, 20) that, for any policy $\pi \in \Pi$,

$$f(\cdot, \pi), g(\cdot, \pi) \in \mathbb{B}_w(\mathbf{X}), \text{ with } \max\{\|f(\cdot, \pi)\|_w, \|g(\cdot, \pi)\|_w\} \leq 2M_\gamma. \quad (21)$$

(ii) For an active region $B \in \mathcal{B}(\mathbf{X})$, the functions $f(\cdot, B)$ and $g(\cdot, B)$ can be evaluated by

$$f(x, B) = \begin{cases} F(x, B) - r(x, 0) - \beta \int F(y, B) Q^0(x, dy), & x \in B \\ r(x, 1) + \beta \int F(y, B) Q^1(x, dy) - F(x, B), & x \in B^c, \end{cases}$$

$$g(x, B) = \begin{cases} G(x, B) - c(x, 0) - \beta \int G(y, B) Q^0(x, dy), & x \in B \\ c(x, 1) + \beta \int G(y, B) Q^1(x, dy) - G(x, B), & x \in B^c; \end{cases}$$

or, in terms of the operator \mathcal{L}^* in (10),

$$f(x, B) = \begin{cases} \mathcal{L}^*F(\cdot, B)(x, 0) - r(x, 0), & x \in B \\ r(x, 1) - \mathcal{L}^*F(\cdot, B)(x, 1), & x \in B^c, \end{cases} \quad (22)$$

$$g(x, B) = \begin{cases} \mathcal{L}^*G(\cdot, B)(x, 0) - c(x, 0), & x \in B \\ c(x, 1) - \mathcal{L}^*G(\cdot, B)(x, 1), & x \in B^c. \end{cases} \quad (23)$$

(iii) For a state x and a policy π , $f(x, \pi)$ and $g(x, \pi)$ can be represented as

$$f(x, \pi) = \Delta_{a=0}^{a=1} \{r(x, a) + \beta F(\kappa^a(x, \cdot), \pi)\} \quad \text{and} \quad g(x, \pi) = \Delta_{a=0}^{a=1} \{c(x, a) + \beta G(\kappa^a(x, \cdot), \pi)\}. \quad (24)$$

If $g(x, B) \neq 0$ for some x and B , we measure the MP of taking the active (vs. the passive) action at $t = 0$ starting from x , provided the B -policy is adopted thereafter, by the *MP metric*

$$m(x, B) \triangleq \frac{f(x, B)}{g(x, B)}. \quad (25)$$

We write as $f(x, z)$, $g(x, z)$, $m(x, z)$ and $f(x, z^-)$, $g(x, z^-)$, $m(x, z^-)$ the marginal metrics for the z -policy and for the z^- -policy, respectively.

We next use the MP metric to define the project's *MP index*.

Definition 2.3 (MP index). Suppose $g(x, x) \neq 0$ for every state x . Then, the project's *MP index* is the function m^* on \mathbf{X} given by

$$m^*(x) \triangleq m(x, x) = \frac{f(x, x)}{g(x, x)}, \quad x \in \mathbf{X}. \quad (26)$$

Henceforth, we will consider the natural extension of the MP index domain from \mathbf{X} to \mathbb{R} , by letting $m^*(z) \triangleq m^*(\ell)$ for $z < \ell$ if $\ell > -\infty$ and $m^*(z) \triangleq m^*(u)$ for $z > u$ if $u < \infty$.

2.6 The main result: A verification theorem for strong threshold-indexability.

This section presents the main result of this paper, a verification theorem providing a set of sufficient conditions for indexability, which, unlike direct application of Definition 2.1, require neither evaluation of the optimal value function V_λ^* of the λ -price problem P_λ in (4) nor knowledge of its properties. Instead, the conditions to be verified here refer to properties of the MP index in (26) and other performance metrics under threshold policies, and hence involve only the *performance analysis* of such policies. The theorem to be presented below shows that, under such conditions, the project is indexable with index m^* .

The sufficient conditions in our verification theorem correspond to the concept of *PCL-indexability*, which we next extend from the discrete state setting where it was introduced and developed in Niño-Mora [34, 35, 36, 37] to the present real state setting.

Definition 2.4 (PCL-indexability). We call a project *PCL-indexable* (with respect to threshold policies) if it satisfies the following *PCL-indexability conditions*:

- (PCLI1) $g(x, z) > 0$ for every state $x \in \mathbf{X}$ and threshold $z \in \overline{\mathbb{R}}$;
- (PCLI2) the MP index m^* is monotone nondecreasing, continuous and bounded below on \mathbf{X} ;
- (PCLI3) for each state x , $F(x, \cdot)$, $G(x, \cdot)$ and m^* are related by

$$F(x, z_2) - F(x, z_1) = \int_{(z_1, z_2]} m^*(z) G(x, dz), \quad -\infty < z_1 < z_2 < \infty. \quad (27)$$

The verification theorem is stated next.

Theorem 2.1. *A PCL-indexable project is strongly threshold-indexable with index m^* .*

Remark 2.5. (i) PCL-indexability conditions (PCLI1, PCLI2) match those given for discrete state projects in Niño-Mora [34, 35, 36], except for the continuity requirement in (PCLI2).

- (ii) Under (PCLI1), the MP metric $m(x, z)$ (see (25)) is well defined for every x and z .
- (iii) Under (PCLI2), we can define $m^*(-\infty) \triangleq \lim_{x \rightarrow -\infty} m^*(x)$, which is finite, and $m^*(\infty) \triangleq \lim_{x \rightarrow \infty} m^*(x)$, which may be finite or infinite.
- (iv) Condition (PCLI3) requires that, for each state x , the reward metric for threshold policies, $F(x, \cdot)$, be an indefinite LS integral of the MP index m^* with respect to the resource metric for threshold policies, $G(x, \cdot)$. The existence and finiteness of such integrals is justified in Lemma 7.4.
- (v) For discrete state projects, (PCLI1) implies (PCLI3). See Niño-Mora [35, The. 6.4(b)] and Niño-Mora [36, Lemma 5.5.(a)].
- (vi) For any initial state distribution $p \in \mathbb{P}_w(\mathbf{X})$, condition (PCLI3) yields (cf. Remark 2.1(iii))

$$F(p, z_2) - F(p, z_1) = \int_{(z_1, z_2]} m^*(z) G(p, dz), \quad -\infty < z_1 < z_2 < \infty. \quad (28)$$

The proof of Theorem 2.1, given in §11, builds on a number of implications of PCL-indexability conditions (PCLI1–PCLI3), which are developed in the following sections.

3 Preliminary results.

This section presents preliminary results shedding light on the PCL-indexability conditions. We start with a characterization of strong threshold-indexability in terms of properties of the project's (Gittins or Whittle) index. This result justifies condition (PCLI2).

Lemma 3.1. *An indexable project with index λ^* is strongly threshold-indexable iff λ^* is monotone nondecreasing and continuous.*

Proof. Proof. Suppose strong threshold-indexability holds. If λ^* were not nondecreasing, there would exist states $x < y$ with $\lambda^*(x) > \lambda^*(y)$. Then for $\lambda \in (\lambda^*(y), \lambda^*(x))$ indexability would imply (see Remark 2.2(i)) that $a = 0$ and $a = 1$ are the only P_λ -optimal actions in y and x , respectively, yielding a contradiction as no threshold policy satisfies such requirements. Hence λ^* must be nondecreasing.

We next show that λ^* must be left-continuous. By monotonicity, λ^* can only have jump discontinuities, so the left limit $\lambda^*(x^-)$ exists at states $x > \ell$ and the right limit $\lambda^*(x^+)$ exists at $x < u$. To argue by contradiction, suppose $\lambda^*(z^-) < \lambda^*(z)$ for some state $z > \ell$. Then for $\lambda \in (\lambda^*(z^-), \lambda^*(z))$ we would have $\lambda^*(x) \leq \lambda^*(z^-) < \lambda < \lambda^*(z)$ for any state $x < z$, whence by indexability (see Remark 2.2(i)) the only P_λ -optimal policy would be the z^- -policy, contradicting strong threshold-indexability.

We next show that λ^* must be right-continuous. If it were $\lambda^*(z) < \lambda^*(z^+)$ for some state $z < u$, then for $\lambda \in (\lambda^*(z), \lambda^*(z^+))$ it would be $\lambda^*(z) < \lambda < \lambda^*(z^+) \leq \lambda^*(y)$ for any state $y > z$, whence by indexability (see Remark 2.2(i)) the only P_λ -optimal policy would be the z -policy, a contradiction.

As for the reverse implication, suppose λ^* is continuous nondecreasing. Fix λ . If $\lambda^*(x) < \lambda$ for every state x then, taking $z = u + 1$, by indexability both the z -policy and the z^- -policy (“always passive”) are P_λ -optimal. If $\lambda^*(x) > \lambda$ for every state x then, taking $z = \ell - 1$, by indexability both the z -policy and the z^- (“always active”) are P_λ -optimal. Otherwise, by continuity of λ^* there is a state z with $\lambda^*(z) = \lambda$ and, by nondecreasingness, $\lambda^*(x) \leq \lambda^*(z) \leq \lambda^*(y)$ for $x < z < y$. By indexability, both the z -policy and the z^- -policy are P_λ -optimal, completing the proof. \square

The following result characterizes via marginal metrics the optimal sets $S_\lambda^{*,a}$ in (16). Let $\lambda \in \mathbb{R}$.

Lemma 3.2. *Let policy π^* be P_λ -optimal. Then*

$$S_\lambda^{*,1} = \{x \in \mathbf{X}: f(x, \pi^*) - \lambda g(x, \pi^*) \geq 0\} \quad \text{and} \quad S_\lambda^{*,0} = \{x \in \mathbf{X}: f(x, \pi^*) - \lambda g(x, \pi^*) \leq 0\}.$$

Proof. Proof. Since π^* is P_λ -optimal, we can reformulate (16) as

$$S_\lambda^{*,a} = \{x \in \mathbf{X}: V_\lambda(x, \langle a, \pi^* \rangle) \geq V_\lambda(x, \langle 1 - a, \pi^* \rangle)\}, \quad a \in \{0, 1\}. \quad (29)$$

Now, it is immediate to verify that

$$\Delta_{a=0}^{a=1} V_\lambda(x, \langle a, \pi^* \rangle) = f(x, \pi^*) - \lambda g(x, \pi^*), \quad (30)$$

which, along with (29), yields the result. \square

The next result applies to projects satisfying (PCLI1). It characterizes via the MP metric $m(x, z)$ (see (25)) the *optimal threshold sets* $Z_\lambda^* \triangleq \{z \in \overline{\mathbb{R}}: \text{the } z\text{-policy is } P_\lambda\text{-optimal}\}$ for $\lambda \in \mathbb{R}$, and the

optimal price sets $\Lambda_z^* \triangleq \{\lambda \in \mathbb{R} : \text{the } z\text{-policy is } P_\lambda\text{-optimal}\}$ for $z \in \overline{\mathbb{R}}$, using the following conditions (where the “sup” and the “inf” over the empty set are taken to be $-\infty$ and ∞ , respectively):

$$\sup_{x \in [\ell, z]} m(x, z) \leq \lambda \leq \inf_{x \in (z, u]} m(x, z). \quad (31)$$

Lemma 3.3. *Let (PCLI1) hold. Then*

- (a) *for each resource price $\lambda \in \mathbb{R}$, $Z_\lambda^* = \{z \in \overline{\mathbb{R}} : (31) \text{ holds}\}$;*
- (b) *for each threshold $z \in \overline{\mathbb{R}}$, $\Lambda_z^* = \{\lambda \in \mathbb{R} : (31) \text{ holds}\}$.*

Proof. Proof. (a) The z -policy is P_λ -optimal iff $V_\lambda(\cdot, z)$ satisfies the BE (13), i.e., iff

$$\bigtriangleup_{a=0}^{a=1} V_\lambda(x, \langle a, z \rangle) \leq 0 \text{ for } x \in [\ell, z] \quad \text{and} \quad \bigtriangleup_{a=0}^{a=1} V_\lambda(x, \langle a, z \rangle) \geq 0 \text{ for } x \in (z, u]. \quad (32)$$

Now, as in (30), it is immediate to verify that

$$\bigtriangleup_{a=0}^{a=1} V_\lambda(x, \langle a, z \rangle) = f(x, z) - \lambda g(x, z), \quad (33)$$

which yields that (31) is a reformulation of (32) under (PCLI1).

- (b) This part follows from (a) and the relation $\Lambda_z^* = \{\lambda \in \mathbb{R} : z \in Z_\lambda^*\}$. □

The next result justifies our choice of the MP index m^* as a candidate for the project’s index.

Proposition 3.1. *Let the project be strongly threshold-indexable with index λ^* . Then*

- (a) $f(x, x) - \lambda^*(x)g(x, x) = 0 = f(x, x^-) - \lambda^*(x)g(x, x^-)$ for every state x ;
- (b) $\lambda^*(x) = m^*(x)$ for any state x with $g(x, x) \neq 0$;
- (c) under (PCLI1), $\lambda^* = m^*$.

Proof. Proof. (a) Taking $\lambda = \lambda^*(x)$, Lemma 3.1 yields that both the x -policy and the x^- -policy are P_λ -optimal, with both the active and the passive actions being P_λ -optimal in x . This implies, by Lemma 3.2, that $f(x, x) - \lambda^*(x)g(x, x) = f(x, x^-) - \lambda^*(x)g(x, x^-) = 0$.

- (b) This part follows from the first identity in (a) and (26).

- (c) This part follows from (b) and (PCLI1). □

4 LP reformulation of the λ -price problem starting from $X_0 \sim p$.

This section lays out further groundwork. It draws on standard results to reformulate the problem

$$P_\lambda(p): \quad \text{maximize}_{\pi \in \Pi} V_\lambda(p, \pi), \quad (34)$$

which is the variant of the λ -price problem P_λ in (4) where the initial state distribution $p \in \mathbb{P}_w(\mathbf{X})$ is fixed, as an infinite-dimensional LP problem over a space of measures.

For an admissible policy π , consider the *state-action occupation measure* μ_p^π defined by

$$\mu_p^\pi(\Gamma) \triangleq \mathbb{E}_p^\pi \left[\sum_{t=0}^{\infty} \beta^t 1_\Gamma(X_t, A_t) \right] = \sum_{t=0}^{\infty} \beta^t \mathbf{P}_p^\pi \{ (X_t, A_t) \in \Gamma \}, \quad \Gamma \in \mathcal{B}(\mathbf{K}). \quad (35)$$

Since the objective of problem $P_\lambda(p)$ can be represented in terms of μ_p^π as

$$V_\lambda(p, \pi) = \int (r - \lambda c) d\mu_p^\pi = \int \{r(y, a) - \lambda c(y, a)\} \mu_p^\pi(d(y, a)),$$

we can reformulate the dynamic optimization problem $P_\lambda(p)$ as the static optimization problem

$$L_\lambda(p): \quad \underset{\mu \in \mathcal{M}_p}{\text{maximize}} \quad \int (r - \lambda c) d\mu, \quad (36)$$

where $\mathcal{M}_p \triangleq \{\mu_p^\pi : \pi \in \Pi\}$ is the achievable region spanned by the μ_p^π as π varies, which is known to be characterized by linear constraints (see Heilmann [15, 16]; Hernández-Lerma and Hernández-Hernández [17], and Hernández-Lerma and Lasserre [18, Sec. 6.3]), as outlined below.

The μ_p^π belong (see Hernández-Lerma and Lasserre [19, Prop. 7.2.2]) to the Banach space $\mathbb{M}_w(\mathbf{K})$ of finite *signed measures* μ on $\mathcal{B}(\mathbf{K})$ with finite w -norm (where $|\mu|$ denotes the *total variation* of μ)

$$\|\mu\|_w \triangleq \sup_{u \in \mathbb{B}_w(\mathbf{K}), \|u\|_w \leq 1} \left| \int u d\mu \right| = \int w d|\mu| = \int w(y) |\mu|(d(y, a)). \quad (37)$$

For each $\mu \in \mathbb{M}_w(\mathbf{K})$ we consider the marginal signed measures on $\mathcal{B}(\mathbf{X})$ defined by $\tilde{\mu}^a(S) \triangleq \mu(S \times \{a\})$ for $a \in \{0, 1\}$, and $\tilde{\mu}(S) \triangleq \mu(S \times \{0, 1\}) = \sum_{a \in \{0, 1\}} \tilde{\mu}^a(S)$, which belong to the Banach space $\mathbb{M}_w(\mathbf{X})$ of finite signed measures ν on $\mathcal{B}(\mathbf{X})$ with finite w -norm

$$\|\nu\|_w \triangleq \sup_{v \in \mathbb{B}_w(\mathbf{X}), \|v\|_w \leq 1} \left| \int v d\nu \right| = \int w d|\nu| = \int w(y) |\nu|(dy). \quad (38)$$

We denote the *marginal state occupation measures* for μ_p^π by $\tilde{\mu}_p^{\pi, a}$ and $\tilde{\mu}_p^\pi$.

Both $(\mathbb{M}_w(\mathbf{K}), \mathbb{B}_w(\mathbf{K}))$ and $(\mathbb{M}_w(\mathbf{X}), \mathbb{B}_w(\mathbf{X}))$ are known to be *dual pairs of vector spaces* (see Hernández-Lerma and Lasserre [19, Sec. 12.2.A]) with respect to the bilinear forms

$$\langle \mu, u \rangle \triangleq \int u d\mu = \int u(y, a) \mu(d(y, a)), \quad \mu \in \mathbb{M}_w(\mathbf{K}), u \in \mathbb{B}_w(\mathbf{K}), \quad (39)$$

$$\langle \nu, v \rangle \triangleq \int v d\nu = \int v(y) \nu(dy), \quad \nu \in \mathbb{M}_w(\mathbf{X}), v \in \mathbb{B}_w(\mathbf{X}). \quad (40)$$

Consider now the linear operators $\mathcal{L}^a: \mathbb{M}_w(\mathbf{K}) \rightarrow \mathbb{M}_w(\mathbf{X})$ for $a \in \{0, 1\}$, mapping μ to

$$\mathcal{L}^a \mu(S) \triangleq \tilde{\mu}^a(S) - \beta \int \kappa^a(y, S) \tilde{\mu}^a(dy), \quad S \in \mathcal{B}(\mathbf{X}), \quad (41)$$

and the linear operator $\mathcal{L} \triangleq \mathcal{L}^0 + \mathcal{L}^1$. Note that

$$\mathcal{L} \mu(S) \triangleq \tilde{\mu}(S) - \beta \int \kappa^a(y, S) \mu(d(y, a)), \quad S \in \mathcal{B}(\mathbf{X}). \quad (42)$$

Such operators are *bounded* under Assumption 2.1(ii.b) (cf. Hernández-Lerma and Lasserre [19, Sec. 7.2.B]).

It is known (cf. Heilmann [15, The. 8] and Hernández-Lerma and Lasserre [18, The. 6.3.7]) that the achievable region \mathcal{M}_p in (36) is a compact and convex region of $\mathbb{M}_w(\mathbf{K})$, which is spanned by stationary randomized policies and is fully characterized by linear constraints as

$$\mathcal{M}_p = \{\mu_p^\pi : \pi \in \Pi^{\text{SR}}\} = \{\mu \in \mathbb{M}_w^+(\mathbf{K}) : \mathcal{L}\mu = p\}, \quad (43)$$

where $\mathbb{M}_w^+(\mathbf{K}) \triangleq \{\mu \in \mathbb{M}_w(\mathbf{K}) : \mu \geq 0\}$ is the *cone of measures* in $\mathbb{M}_w(\mathbf{K})$.

We can thus explicitly formulate problem $L_\lambda(p)$ in (36) as the infinite-dimensional LP problem

$$\begin{aligned} L_\lambda(p): \quad & \text{maximize } \langle \mu, r - \lambda c \rangle \\ & \text{subject to: } \mathcal{L}\mu = p, \mu \in \mathbb{M}_w^+(\mathbf{K}). \end{aligned} \quad (44)$$

Further, using that $\mathcal{L} = \sum_a \mathcal{L}^a$, (41), and decoupling μ into $\tilde{\mu}^0$ and $\tilde{\mu}^1$, yields the following equivalent LP reformulation of problem $L_\lambda(p)$ in the variables $\tilde{\mu}^0$ and $\tilde{\mu}^1$:

$$\begin{aligned} L_\lambda(p): \quad & \text{maximize } \sum_{a \in \{0,1\}} \langle \tilde{\mu}^a, r(\cdot, a) - \lambda c(\cdot, a) \rangle \\ & \text{subject to: } \tilde{\mu}^0, \tilde{\mu}^1 \in \mathbb{M}_w^+(\mathbf{X}) \\ & \sum_{a \in \{0,1\}} \left\{ \tilde{\mu}^a - \beta \int \kappa^a(y, \cdot) \tilde{\mu}^a(dy) \right\} = p, \end{aligned} \quad (45)$$

where $\mathbb{M}_w^+(\mathbf{X}) \triangleq \{\nu \in \mathbb{M}_w(\mathbf{X}) : \nu \geq 0\}$ is the cone of measures in $\mathbb{M}_w(\mathbf{X})$.

In particular, the above results apply when the initial state $X_0 = x$ is fixed, so $p = \delta_x$, the Dirac measure concentrated at x . In the sequel, we will thus refer to the problem $P_\lambda(x)$, given by

$$P_\lambda(x): \quad \text{maximize } V_\lambda(x, \pi), \quad \pi \in \Pi \quad (46)$$

and to its LP reformulation $L_\lambda(x)$, given by

$$\begin{aligned} L_\lambda(x): \quad & \text{maximize } \sum_{a \in \{0,1\}} \langle \tilde{\mu}^a, r(\cdot, a) - \lambda c(\cdot, a) \rangle \\ & \text{subject to: } \tilde{\mu}^0, \tilde{\mu}^1 \in \mathbb{M}_w^+(\mathbf{X}) \\ & \sum_{a \in \{0,1\}} \left\{ \tilde{\mu}^a - \beta \int \kappa^a(y, \cdot) \tilde{\mu}^a(dy) \right\} = \delta_x. \end{aligned} \quad (47)$$

5 Decomposition of performance metrics.

This section presents a decomposition of the project's performance metrics, which will be used in subsequent analyses.

We first need a preliminary result. Note that the linear operator \mathcal{L}^* in (10) is (cf. Hernández-Lerma and Lasserre [18, p. 139]) the *adjoint* of \mathcal{L} in (42), as it satisfies the *basic adjoint relation*

$$\langle \mathcal{L}\mu, v \rangle = \langle \mu, \mathcal{L}^*v \rangle, \quad \mu \in \mathbb{M}_w(\mathbf{K}), v \in \mathbb{B}_w(\mathbf{X}), \quad (48)$$

or, in terms of the operators \mathcal{L}^a in (41),

$$\sum_{a \in \{0,1\}} \langle \mathcal{L}^a \mu, v \rangle = \langle \mu, \mathcal{L}^*v \rangle. \quad (49)$$

The next result (which the author has not found in the literature) gives a decomposition of (49).

Lemma 5.1. For any $a \in \{0, 1\}$, $\mu \in \mathbb{M}_w(\mathbf{K})$ and $v \in \mathbb{B}_w(\mathbf{X})$,

$$\langle \mathcal{L}^a \mu, v \rangle = \langle \tilde{\mu}^a, \mathcal{L}^* v(\cdot, a) \rangle.$$

Proof. Proof. Let $a \in \{0, 1\}$ and $\mu \in \mathbb{M}_w(\mathbf{K})$. For any $S \in \mathcal{B}(\mathbf{X})$, we have

$$\begin{aligned} \langle \mathcal{L}^a \mu, 1_S \rangle &= \mathcal{L}^a \mu(S) = \int \left\{ 1_S(x) - \beta \int 1_S(y) Q^a(x, dy) \right\} \tilde{\mu}^a(dx) \\ &= \int \mathcal{L}^* 1_S(x, a) \tilde{\mu}^a(dx) = \langle \tilde{\mu}^a, \mathcal{L}^* 1_S(\cdot, a) \rangle, \end{aligned}$$

where we have used in turn (40), (41), (10), and (39). By standard arguments, it follows from the above that $\langle \mathcal{L}^a \mu, v \rangle = \langle \tilde{\mu}^a, \mathcal{L}^* v(\cdot, a) \rangle$ for any $v \in \mathbb{B}_w(\mathbf{X})$. \square

We next give the main result of this section, consisting of two parts. Part (a) decomposes the reward metric $F(p, \pi)$ as the sum of $F(p, B)$ for a given active region B and a linear combination of $\tilde{\mu}_p^{\pi, a}$, with coefficients given by marginal rewards $f(y, B)$. Part (b) similarly decomposes the resource metric $G(p, \pi)$. This result extends to the real state setting corresponding results for the finite and countably-infinite state cases given in Niño-Mora [34, The. 3], Niño-Mora [35, Prop. 6.1] and Niño-Mora [36, Lemma 5.4]. Let $p \in \mathbb{P}_w(\mathbf{X})$.

Lemma 5.2. For any policy $\pi \in \Pi$ and active region $B \in \mathcal{B}(\mathbf{X})$,

$$(a) \quad F(p, \pi) = F(p, B) - \int_B f(y, B) \tilde{\mu}_p^{\pi, 0}(dy) + \int_{B^c} f(y, B) \tilde{\mu}_p^{\pi, 1}(dy), \text{ viz.,}$$

$$F(p, \pi) = F(p, B) + \mathbb{E}_p^\pi \left[\sum_{t=0}^{\infty} \beta^t \{A_t - 1_B(X_t)\} f(X_t, B) \right]; \quad (50)$$

$$(b) \quad G(p, \pi) = G(p, B) - \int_B g(y, B) \tilde{\mu}_p^{\pi, 0}(dy) + \int_{B^c} g(y, B) \tilde{\mu}_p^{\pi, 1}(dy), \text{ viz.,}$$

$$G(p, \pi) = G(p, B) + \mathbb{E}_p^\pi \left[\sum_{t=0}^{\infty} \beta^t \{A_t - 1_B(X_t)\} g(X_t, B) \right]. \quad (51)$$

Proof. Proof. (a) Using in turn Lemma 5.1, (40), (22) and (11), we can write

$$\begin{aligned} \langle \mathcal{L}^0 \mu_p^\pi, F(\cdot, B) \rangle &= \langle \tilde{\mu}_p^{\pi, 0}, \mathcal{L}^* F(\cdot, B)(\cdot, 0) \rangle = \int \mathcal{L}^* F(\cdot, B)(y, 0) \tilde{\mu}_p^{\pi, 0}(dy) \\ &= \int_B \mathcal{L}^* F(\cdot, B)(y, 0) \tilde{\mu}_p^{\pi, 0}(dy) + \int_{B^c} \mathcal{L}^* F(\cdot, B)(y, 0) \tilde{\mu}_p^{\pi, 0}(dy) \\ &= \int_B \{r(y, 0) + f(y, B)\} \tilde{\mu}_p^{\pi, 0}(dy) + \int_{B^c} r(y, 0) \tilde{\mu}_p^{\pi, 0}(dy) \\ &= \int r(y, 0) \tilde{\mu}_p^{\pi, 0}(dy) + \int_B f(y, B) \tilde{\mu}_p^{\pi, 0}(dy), \end{aligned}$$

and

$$\begin{aligned} \langle \mathcal{L}^1 \mu_p^\pi, F(\cdot, B) \rangle &= \langle \tilde{\mu}_p^{\pi, 1}, \mathcal{L}^* F(\cdot, B)(\cdot, 1) \rangle = \int \mathcal{L}^* F(\cdot, B)(y, 1) \tilde{\mu}_p^{\pi, 1}(dy) \\ &= \int_B \mathcal{L}^* F(\cdot, B)(y, 1) \tilde{\mu}_p^{\pi, 1}(dy) + \int_{B^c} \mathcal{L}^* F(\cdot, B)(y, 1) \tilde{\mu}_p^{\pi, 1}(dy) \\ &= \int_B r(y, 1) \tilde{\mu}_p^{\pi, 1}(dy) + \int_{B^c} \{r(y, 1) - f(y, B)\} \tilde{\mu}_p^{\pi, 1}(dy) \\ &= \int r(y, 1) \tilde{\mu}_p^{\pi, 1}(dy) - \int_{B^c} f(y, B) \tilde{\mu}_p^{\pi, 1}(dy). \end{aligned}$$

Further, using that $\mathcal{L} = \sum_{a \in \{0,1\}} \mathcal{L}^a$, (43) and (40) we obtain

$$\sum_{a \in \{0,1\}} \langle \mathcal{L}^a \mu_p^\pi, F(\cdot, B) \rangle = \langle \mathcal{L} \mu_p^\pi, F(\cdot, B) \rangle = \langle p, F(\cdot, B) \rangle = F(p, B).$$

The result now follows by combining the above identities along with

$$F(p, \pi) = \int r d\mu_p^\pi = \sum_{a \in \{0,1\}} \int r(y, a) \tilde{\mu}_p^{\pi, a}(dy).$$

As for (50), it follows by reformulating the decomposition as

$$F(p, \pi) + \mathbb{E}_p^\pi \left[\sum_{t=0}^{\infty} \beta^t 1_{B \times \{0\}}(X_t, A_t) f(X_t, B) \right] = F(p, B) + \mathbb{E}_p^\pi \left[\sum_{t=0}^{\infty} \beta^t 1_{B^c \times \{1\}}(X_t, A_t) f(X_t, B) \right],$$

and then rearranging and simplifying the latter identity as stated.

Part (b) follows along the same lines as part (a). \square

We will use Lemma 5.2 as a powerful tool to generate useful identities. Thus, we use it next to reformulate the objective of problem $P_\lambda(x)$ in (46) in a form that will be more convenient for our analyses, using only the resource metric $G(x, \pi)$ and the passive state occupation measure $\tilde{\mu}_x^{\pi, 0}$. Taking in Lemma 5.2(a) the active region $B = \mathbf{X}$ for the ℓ^- -policy (“always active”) gives

$$F(x, \pi) = F(x, \ell^-) - \int f(y, \ell^-) \tilde{\mu}_x^{\pi, 0}(dy), \quad (52)$$

whence problem $P_\lambda(x)$ is equivalent to

$$\tilde{P}_\lambda(x): \quad \underset{\pi \in \Pi}{\text{minimize}} \quad \int f(y, \ell^-) \tilde{\mu}_x^{\pi, 0}(dy) + \lambda G(x, \pi). \quad (53)$$

We will denote the objective value of policy π in problem $\tilde{P}_\lambda(x)$ by $\tilde{V}_\lambda(x, \pi) \triangleq F(x, \ell^-) - V_\lambda(x, \pi)$. The optimal objective value of problem $\tilde{P}_\lambda(x)$ is hence $\tilde{V}_\lambda^*(x) \triangleq F(x, \ell^-) - V_\lambda^*(x)$.

Note that the LP reformulation $L_\lambda(x)$ of problem $P_\lambda(x)$ given in (47) immediately yields the following LP reformulation of the equivalent problem $\tilde{P}_\lambda(x)$:

$$\begin{aligned} \tilde{L}_\lambda(x): \quad & \text{minimize} \quad \int f(y, \ell^-) \tilde{\mu}^0(dy) + \lambda \gamma \\ & \text{subject to: } \tilde{\mu}^0, \tilde{\mu}^1 \in \mathbb{M}_w^+(\mathbf{X}), \gamma \in \mathbb{R} \\ & \sum_{a \in \{0,1\}} \left\{ \tilde{\mu}^a - \beta \int \kappa^a(y, \cdot) \tilde{\mu}^a(dy) \right\} = \delta_x \\ & \gamma - \sum_{a \in \{0,1\}} \int c(y, a) \tilde{\mu}^a(dy) = 0. \end{aligned} \quad (54)$$

6 Partial conservation laws and a relaxation of $\tilde{L}_\lambda(x)$.

We will find it convenient to work with a reformulation of $\tilde{L}_\lambda(x)$ in the variables $\tilde{\mu}^0$ and γ only. Rather than projecting out $\tilde{\mu}^1$ in (54), we will construct a *relaxation* of $\tilde{L}_\lambda(x)$ in $\tilde{\mu}^0$ and γ , by deriving a collection of linear constraints satisfied by the metrics $\tilde{\mu}_x^{\pi,0}$ and $G(x, \pi)$. For such a purpose, we will use the identities in Lemma 5.2(b) using active regions B for threshold policies. The resulting constraints take the form of *partial conservation laws* (PCLs). The following definition extends the concept of PCLs from the discrete state setting (see §1.2) to the present real state setting.

Definition 6.1 (PCLs). We say the project's performance metrics $\tilde{\mu}_x^{\pi,0}, G(x, \pi)$ satisfy PCLs (with respect to threshold policies) if the following holds: for any state x and admissible policy π ,

- (a) $\int g(y, \ell^-) \tilde{\mu}_x^{\pi,0}(dy) + G(x, \pi) = G(x, \ell^-);$
 - (b.1) $G(x, \pi) \geq G(x, u);$
 - (c.1) $\int_{(b,u]} g(y, b) \tilde{\mu}_x^{\pi,0}(dy) + G(x, \pi) \geq G(x, b), \quad \ell \leq b < u;$
 - (d.1) $\int_{(b,u]} g(y, b) \tilde{\mu}_x^{\pi,0}(dy) \geq 0, \quad \ell \leq b < u;$
- and, furthermore, for π stationary deterministic,
- (b.2) equality holds in (b.1) for every state x iff π selects the passive action on \mathbf{X} ;
 - (c.2) equality holds in (c.1) for every state x iff π selects the passive action on $[\ell, b]$;
 - (d.2) equality holds in (d.1) for every state x iff π selects the active action on $(b, u]$.

We call such conservation laws “partial” because they apply only to the family of active regions of the form $B = (b, u]$ for thresholds b , rather than to all active regions $B \in \mathcal{B}(\mathbf{X})$.

- Remark 6.1.* (i) PCLs consist of: (1) the identity in (a) and the inequalities in (b.1), (c.1) and (d.1), which hold for any admissible policy; and (2) the identities in (b.2), (c.2) and (d.2), which hold for the stationary deterministic policies satisfying the stated properties.
- (ii) PCLs characterize the optimal stationary deterministic policies for each of the following problems, which are special cases of (53): $\min_\pi \int g(y, \ell^-) \tilde{\mu}_x^{\pi,0}(dy) + G(x, \pi)$; $\min_\pi G(x, \pi)$;

$$\min_\pi \int_{(b,u]} g(y, b) \tilde{\mu}_x^{\pi,0}(dy) + G(x, \pi), \quad \ell \leq b < u;$$

and $\min_\pi \int_{(b,u]} g(y, b) \tilde{\mu}_x^{\pi,0}(dy)$, for $\ell \leq b < u$.

The next result shows that condition (PCLI1) in Definition 2.4 implies satisfaction of PCLs.

Lemma 6.1. *Under (PCLI1), the performance metrics $\tilde{\mu}_x^{\pi,0}, G(x, \pi)$ satisfy PCLs.*

Proof. Proof. We show in turn that each part of Definition 6.1 holds under (PCLI1).

- (a) This identity follows by taking $B = \mathbf{X}$ in Lemma 5.2(b).

(b.1) Taking the active region $B = (u, u] = \emptyset$ in Lemma 5.2(b) and using (PCLI1) yields

$$G(x, \pi) = G(x, u) + \int g(y, u) \tilde{\mu}_x^{\pi,1}(dy) \geq G(x, u). \quad (55)$$

(b.2) Let $\pi \in \Pi^{\text{SD}}$. If π rests the project on \mathbf{X} then $\tilde{\mu}_x^{\pi,1}(\mathbf{X}) = 0$ and (55) yields $G(x, \pi) = G(x, u)$. Conversely, $G(x, \pi) = G(x, u)$ for every x , (55) and (PCLI1) yield

$$0 = \int g(y, u) \tilde{\mu}_x^{\pi,1}(dy) = \mathbb{E}_x^\pi \left[\sum_{t=0}^{\infty} \beta^t g(X_t, u) A_t \right] \geq g(x, u) \mathbb{P}_x^\pi \{A_0 = 1\} \geq 0, \quad x \in \mathbf{X},$$

whence $\mathbb{P}_x^\pi \{A_0 = 1\} = 0$ for every x , i.e., π is passive on \mathbf{X} .

(c.1) Taking the active region $B = (b, u]$ in Lemma 5.2(b) and using (PCLI1) yields

$$G(x, \pi) + \int_{(b,u]} g(y, b) \tilde{\mu}_x^{\pi,0}(dy) = G(x, b) + \int_{[\ell,b]} g(y, b) \tilde{\mu}_x^{\pi,1}(dy) \geq G(x, b). \quad (56)$$

(c.2) Let $\pi \in \Pi^{\text{SD}}$. If π is passive on $[\ell, b]$ then $\tilde{\mu}_x^{\pi,1}[\ell, b] = 0$, $\int_{[\ell,b]} g(y, b) \tilde{\mu}_x^{\pi,1}(dy) = 0$, and

$$G(x, \pi) + \int_{(b,u]} g(y, b) \tilde{\mu}_x^{\pi,0}(dy) = G(x, b).$$

Conversely, if the latter identity holds for every x , then we have, by (PCLI1),

$$0 = \int_{[\ell,b]} g(y, b) \tilde{\mu}_x^{\pi,1}(dy) \geq g(x, b) \mathbb{P}_x^\pi \{A_0 = 1\} \geq 0, \quad x \in [\ell, b],$$

whence $\mathbb{P}_x^\pi \{A_0 = 1\} = 0$ for every state $x \in [\ell, b]$, i.e., π rests the project on $[\ell, b]$.

Parts (d.1) and (d.2) follow along the same lines as the previous parts. \square

The next result draws on the above to formulate a relaxation of problem $\tilde{L}_\lambda(x)$ in (54).

Lemma 6.2. *Under (PCLI1), problem $R_\lambda(x)$ below is a relaxation of $\tilde{L}_\lambda(x)$, in that if $(\tilde{\mu}^0, \tilde{\mu}^1, \gamma)$ is an $\tilde{L}_\lambda(x)$ -feasible solution then $(\tilde{\mu}^0, \gamma)$ is $R_\lambda(x)$ -feasible.*

$$\begin{aligned} R_\lambda(x): \quad & \text{minimize} \quad \int f(y, \ell^-) \tilde{\mu}^0(dy) + \lambda \gamma \\ & \text{subject to: } \tilde{\mu}^0 \in \mathbb{M}_w^+(\mathbf{X}), \gamma \in \mathbb{R} \\ \psi(db): \quad & \int_{(b,u]} g(y, b) \tilde{\mu}^0(dy) + \gamma \geq G(x, b), \quad b \in [\ell, u) \\ \eta: \quad & \int g(y, \ell^-) \tilde{\mu}^0(dy) + \gamma = G(x, \ell^-) \\ \xi: \quad & \gamma \geq G(x, u). \end{aligned} \quad (57)$$

Proof. Proof. The result follows from (52), Lemma 6.1 and (43). \square

We will denote by $\text{obj}(\tilde{\mu}^0, \gamma; R_\lambda)$ the objective value in (57) of an $R_\lambda(x)$ -feasible solution $(\tilde{\mu}^0, \gamma)$, and by $\text{val}(R_\lambda(x))$ the optimal objective value of problem $R_\lambda(x)$.

Remark 6.2. (i) The optimal values of $R_\lambda(x)$ and $\tilde{L}_\lambda(x)$ satisfy $\text{val}(R_\lambda(x)) \leq \tilde{V}_\lambda^*(x)$.

- (ii) The formulation of $R_\lambda(x)$ does not include the constraints corresponding to the PCLs in Definition 6.1(d.1), because they are redundant under (PCLI1).
- (iii) To prepare the ground for later analyses, we have indicated in (57) the dual variable corresponding to each constraint of $R_\lambda(x)$: scalars η, ξ and a real-valued function ψ on \mathbf{X} .

Under PCL-indexability, we will solve the relaxed problem $R_\lambda(x)$ in closed form using LP duality, showing that its optimal value equals that of $\tilde{L}_\lambda(x)$. For such a purpose, we need further preliminary results, which are developed in the following section.

7 Performance metrics as functions of the threshold.

This section presents properties of the performance metrics as functions of the threshold, which are used in the subsequent analyses. Throughout, let $p \in \mathbb{P}_w(\mathbf{X})$ and $x \in \mathbf{X}$.

7.1 Càdlàg property.

We show, in Lemma 7.2, that $F(p, z)$, $G(p, z)$, $f(x, z)$ and $g(x, z)$ are càdlàg functions of the threshold z . We need the following preliminary result, referring to the natural extension to \mathbb{R} of the state occupation measures $\tilde{\mu}_p^\pi$, viz., $\tilde{\mu}_p^\pi(S) \triangleq \tilde{\mu}_p^\pi(S \cap \mathbf{X})$ for $S \in \mathcal{B}(\mathbb{R})$.

Lemma 7.1. *For any thresholds $-\infty \leq z_1 < z_2 \leq \infty$,*

- (a) $F(p, z_1) - F(p, z_2) = \int_{(z_1, z_2]} f(y, z_2) \tilde{\mu}_p^{z_1}(dy) = \int_{(z_1, z_2]} f(y, z_1) \tilde{\mu}_p^{z_2}(dy);$
- (b) $G(p, z_1) - G(p, z_2) = \int_{(z_1, z_2]} g(y, z_2) \tilde{\mu}_p^{z_1}(dy) = \int_{(z_1, z_2]} g(y, z_1) \tilde{\mu}_p^{z_2}(dy);$
- (c) $F(p, z_1) - F(p, z_2^-) = \int_{(z_1, z_2)} f(y, z_2^-) \tilde{\mu}_p^{z_1}(dy) = \int_{(z_1, z_2)} f(y, z_1) \tilde{\mu}_p^{z_2^-}(dy);$
- (d) $G(p, z_1) - G(p, z_2^-) = \int_{(z_1, z_2)} g(y, z_2^-) \tilde{\mu}_p^{z_1}(dy) = \int_{(z_1, z_2)} g(y, z_1) \tilde{\mu}_p^{z_2^-}(dy).$
- (e) $F(p, z_1^-) - F(p, z_2) = \int_{[z_1, z_2]} f(y, z_2) \tilde{\mu}_p^{z_1^-}(dy) = \int_{[z_1, z_2]} f(y, z_1^-) \tilde{\mu}_p^{z_2}(dy);$
- (f) $G(p, z_1^-) - G(p, z_2) = \int_{[z_1, z_2]} g(y, z_2) \tilde{\mu}_p^{z_1^-}(dy) = \int_{[z_1, z_2]} g(y, z_1^-) \tilde{\mu}_p^{z_2}(dy).$

Proof. Proof. (a, b) For the first identities take $\pi = z_1$ and $B = (z_2, u]$ in Lemma 5.2(a, b), using that $\tilde{\mu}_p^{z_1, 0}(z_1, u] = \tilde{\mu}_p^{z_1, 1}[\ell, z_1] = 0$ and $\tilde{\mu}_p^{z_1, 1} = \tilde{\mu}_p^{z_1}$ on $\mathcal{B}(z_1, z_2]$. For the second identities take $\pi = z_2$ and $B = (z_1, u]$, using that $\tilde{\mu}_p^{z_2, 0}(z_2, u] = \tilde{\mu}_p^{z_2, 1}[\ell, z_1] = 0$ and $\tilde{\mu}_p^{z_2, 0} = \tilde{\mu}_p^{z_2}$ on $\mathcal{B}(z_1, z_2]$.

(c, d) For the first identities take $\pi = z_1$ and $B = [z_2, u]$ in Lemma 5.2(a, b), using that $\tilde{\mu}_p^{z_1, 0}[z_2, u] = \tilde{\mu}_p^{z_1, 1}[\ell, z_1] = 0$ and $\tilde{\mu}_p^{z_1, 1} = \tilde{\mu}_p^{z_1}$ on $\mathcal{B}(z_1, z_2)$. For the second identities take $\pi = z_2^-$ and $B = (z_1, u]$, using that $\tilde{\mu}_p^{z_2^-, 0}[z_2, u] = \tilde{\mu}_p^{z_2^-, 1}[\ell, z_1] = 0$ and $\tilde{\mu}_p^{z_2^-, 0} = \tilde{\mu}_p^{z_2^-}$ on $\mathcal{B}(z_1, z_2)$.

(e, f) For the first identities take $\pi = z_1^-$ and $B = (z_2, u]$ in Lemma 5.2(a, b), using that $\tilde{\mu}_p^{z_1^-, 0}(z_2, u] = \tilde{\mu}_p^{z_1^-, 1}(\bar{\mathbf{X}}_{z_1}^c) = 0$ and $\tilde{\mu}_p^{z_1^-, 1} = \tilde{\mu}_p^{z_1^-}$ on $\mathcal{B}[z_1, z_2]$. For the second identities take $\pi = z_2$ and $B = [z_1, u]$, using that $\tilde{\mu}_p^{z_2, 0}(z_2, u] = \tilde{\mu}_p^{z_2, 1}[\ell, z_1] = 0$ and $\tilde{\mu}_p^{z_2, 0} = \tilde{\mu}_p^{z_2}$ on $\mathcal{B}[z_1, z_2]$. \square

Recall from §2.4 that the notation $F(p, z^-)$ (for this and other performance metrics) refers to the metric $F(p, \pi)$ under the z^- -policy. The next result ensures the consistency of such notation with the standard one where $F(p, z^-)$ denotes the left limit at z of function $F(p, \cdot)$. The result further shows the consistency of the notation $F(p, -\infty)$ and $F(p, \infty)$, where the second argument refers to the threshold policies $-\infty$ (“always active”) and ∞ (“never active”) with the standard notation where $F(p, -\infty)$ and $F(p, \infty)$ are the limits of $F(p, z)$ as $z \rightarrow -\infty$ and as $z \rightarrow \infty$.

Lemma 7.2.

- (a) $F(p, \cdot)$, $G(p, \cdot)$, $f(x, \cdot)$ and $g(x, \cdot)$ are bounded càdlàg functions with left limits at $z \in \mathbb{R}$ given by $F(p, z^-)$, $G(p, z^-)$, $f(x, z^-)$ and $g(x, z^-)$;
- (b) $F(p, z)$, $G(p, z)$, $f(x, z)$ and $g(x, z)$ converge as $z \rightarrow -\infty$ to $F(p, -\infty)$, $G(p, -\infty)$, $f(x, -\infty)$ and $g(x, -\infty)$;
- (c) $F(p, z)$, $G(p, z)$, $f(x, z)$ and $g(x, z)$ converge as $z \rightarrow \infty$ to $F(p, \infty)$, $G(p, \infty)$, $f(x, \infty)$ and $g(x, \infty)$.

Proof. Proof. (a) We first show that $F(p, \cdot)$ is right-continuous at z . From Lemma 7.1(a) we obtain (note that M_γ is as in (7)), for $\delta > 0$,

$$|F(p, z + \delta) - F(p, z)| \leq \int_{(z, z+\delta]} |f|(y, z + \delta) \tilde{\mu}_p^z(dy) \leq 2M_\gamma \int_{(z, z+\delta]} w(y) \tilde{\mu}_p^z(dy) \rightarrow 0 \text{ as } \delta \searrow 0,$$

where the second inequality follows from (21), and the limit follows by the dominated convergence theorem, since $\tilde{\mu}_p^z \in \mathbb{M}_w(\mathbf{X})$ (see (38)). The result for $G(p, \cdot)$ follows similarly. Furthermore,

$$\begin{aligned} \lim_{\delta \searrow 0} f(x, z + \delta) &= \lim_{\delta \searrow 0} \Delta_{a=0}^{a=1} \left\{ r(x, a) + \beta \int F(y, z + \delta) Q^a(x, dy) \right\} \\ &= \Delta_{a=0}^{a=1} \left\{ r(x, a) + \beta \int F(y, z) Q^a(x, dy) \right\} = f(x, z), \end{aligned}$$

where the interchange of limit and integral is justified by the dominated convergence theorem, using (7) and part (a). The result for $g(x, \cdot)$ follows similarly.

Consider now the left limits. From Lemma 7.1(c) and arguing as above we obtain

$$|F(p, z - \delta) - F(p, z^-)| \leq \int_{(z-\delta, z)} |f|(y, z - \delta) \tilde{\mu}_p^{z^-}(dy) \leq 2M_\gamma \int_{(z-\delta, z)} w(y) \tilde{\mu}_p^{z^-}(dy) \rightarrow 0 \text{ as } \delta \searrow 0,$$

and similarly for the left limits of $G(p, \cdot)$, $f(x, \cdot)$ and $g(x, \cdot)$ at z .

(b) The result is trivial if ℓ is finite, so suppose $\ell = -\infty$. For any finite z , taking $z_1 = -\infty$ and $z_2 = z$ in the first identity in Lemma 7.1(a) yields

$$F(p, -\infty) = F(p, z) + \int_{(-\infty, z]} f(y, z) \tilde{\mu}_p^{-\infty}(dy). \quad (58)$$

The result $\lim_{z \rightarrow -\infty} F(p, z) = F(p, -\infty)$ now follows from (58) by the dominated convergence theorem, arguing as in part (a). The results for $G(p, -\infty)$, $f(x, -\infty)$ and $g(x, -\infty)$ follow similarly.

(c) The result is trivial if u is finite, so suppose $u = \infty$. For any finite z , taking $z_1 = z$ and $z_2 = \infty$ in the second identity in Lemma 7.1(a) yields

$$F(p, \infty) = F(p, z) - \int_{(z, \infty)} f(y, z) \tilde{\mu}_p^\infty(dy). \quad (59)$$

The result $\lim_{z \rightarrow \infty} F(p, z) = F(p, \infty)$ follows from (59) by the dominated convergence theorem, arguing as in part (a). The results for $G(p, \infty)$, $f(x, \infty)$ and $g(x, \infty)$ follow similarly. \square

7.2 Monotonicity of resource metric.

The following result shows that, under condition (PCLI1), the resource metric $G(p, z)$ is, as intuition would suggest, monotone nonincreasing in the threshold variable z . Note that in part (b) the term “decreasing” is used in the strict sense.

Lemma 7.3. *Let (PCLI1) hold. Then*

- (a) $G(p, z)$ is nonincreasing in z on \mathbb{R} ;
- (b) if p has full support X then $G(p, z)$ is decreasing in z on X .

Proof. Proof. (a) For finite $z_1 < z_2$, Lemma 7.1(b) and (PCLI1) yield $G(p, z_1) \geq G(p, z_2)$.

(b) Lemma 7.1(b) further yields that, for any finite thresholds $\ell \leq z_1 < z_2 \leq u$,

$$G(p, z_1) - G(p, z_2) = \int_{(z_1, z_2]} g(y, z_2) \tilde{\mu}_p^{z_1}(dy) > 0, \quad (60)$$

where the inequality follows from (PCLI1) and $\tilde{\mu}_p^{z_1}(z_1, z_2] \geq p(z_1, z_2] > 0$. \square

We next establish the existence and finiteness of the LS integrals in condition (PCLI3).

Lemma 7.4. *Under (PCLI1, PCLI2), $\int_{(z_1, z_2]} m^*(z) G(p, dz)$ exists and is finite for finite $z_1 < z_2$.*

Proof. Proof. By Lemmas 7.2 and 7.3(a), $G(p, \cdot)$ is bounded, càdlàg and nonincreasing. Since m^* is continuous, $\int_{(z_1, z_2]} m^*(z) G(p, dz)$ exists and is finite by standard results on LS integration. \square

7.3 Analysis of discontinuities.

Being càdlàg, the functions $F(p, \cdot)$, $G(p, \cdot)$, $f(x, \cdot)$ and $g(x, \cdot)$ have at most countably many discontinuities, which are of jump type. This section analyzes such discontinuities, starting with the following result, which gives simple formulae for their jumps.

Lemma 7.5. *For any threshold $z \in \mathsf{X}$,*

- (a) $\Delta_2 F(p, z) = -f(z, z^-) \tilde{\mu}_p^z\{z\} = -f(z, z) \tilde{\mu}_p^{z^-}\{z\}$;
- (b) $\Delta_2 G(p, z) = -g(z, z^-) \tilde{\mu}_p^z\{z\} = -g(z, z) \tilde{\mu}_p^{z^-}\{z\}$.

Proof. Proof. (a) Taking $\pi = z$ and $B = [z, u]$ in Lemma 5.2(a) gives

$$F(p, z) + \int_{[z, u]} f(y, z^-) \tilde{\mu}_p^{z, 0}(dy) = F(p, z^-) + \int_{[\ell, z]} f(y, z^-) \tilde{\mu}_p^{z, 1}(dy),$$

so $\Delta_2 F(p, z) + f(z, z^-) \tilde{\mu}_p^z \{z\} = 0$, as $\tilde{\mu}_p^{z, 0} \{z\} = \tilde{\mu}_p^z \{z\}$ and $\tilde{\mu}_p^{z, 0}(z, u] = \tilde{\mu}_p^{z, 1}[\ell, z) = 0$.

Further, taking $\pi = z^-$ and $B = (z, u]$ in Lemma 5.2(a) gives

$$F(p, z^-) + \int_{(z, u]} f(y, z) \tilde{\mu}_p^{z^-, 0}(dy) = F(p, z) + \int_{[\ell, z]} f(y, z) \tilde{\mu}_p^{z^-, 1}(dy),$$

so $\Delta_2 F(p, z) + f(z, z) \tilde{\mu}_p^{z^-} \{z\} = 0$, as $\tilde{\mu}_p^{z^-, 1} \{z\} = \tilde{\mu}_p^{z^-} \{z\}$ and $\tilde{\mu}_p^{z^-, 0}(z, u] = \tilde{\mu}_p^{z^-, 1}[\ell, z) = 0$.

Part (b) follows along the same lines as part (a). \square

Recall that PCL-indexability condition (PCLI1) requires $g(x, z)$ to be positive for every state x and threshold z . The following result shows that it further implies $g(x, x^-) > 0$.

Lemma 7.6. *Under (PCLI1), $g(x, x^-) > 0$.*

Proof. Proof. Taking $z = x$ in Lemma 7.5(b) and using (PCLI1) and $\tilde{\mu}_x^{x^-} \{x\} \geq 1$ gives

$$g(x, x^-) \tilde{\mu}_x^x \{x\} = g(x, x) \tilde{\mu}_x^{x^-} \{x\} \geq g(x, x) > 0,$$

whence $g(x, x^-) > 0$. \square

The next result characterizes the discontinuities of $G(p, \cdot)$ in terms of $T_z \triangleq \min\{t \geq 0 : X_t = z\}$, the *first hitting time to z* of $\{X_t\}_{t=0}^\infty$ under the z -policy, with $T_z \triangleq \infty$ if $\{X_t\}_{t=0}^\infty$ never hits z .

Lemma 7.7. *Under (PCLI1), $G(p, \cdot)$ is discontinuous at z iff $\mathbf{P}_p^z\{T_z < \infty\} > 0$.*

Proof. Proof. It follows from Lemmas 7.5(b) and 7.6 that $G(p, \cdot)$ is discontinuous at z iff $\tilde{\mu}_p^z \{z\} > 0$, which happens iff $\mathbf{P}_p^z\{T_z < \infty\} > 0$. \square

The following result relates the jumps of $F(p, \cdot)$ to those of $G(p, \cdot)$, gives a corresponding result for $f(x, \cdot)$ and $g(x, \cdot)$, and ensures that the MP metric $m(x, z)$ (see (25)) is continuous at $z = x$.

Lemma 7.8. *Let (PCLI1) hold. Then, for any threshold $z \in \mathbf{X}$,*

- (a) $\Delta_2 F(p, z) = m^*(z) \Delta_2 G(p, z)$;
- (b) $\Delta_2 f(x, z) = m^*(z) \Delta_2 g(x, z)$;
- (c) $m(x, \cdot)$ is continuous at $z = x$, viz., $m(x, x^-) = m(x, x) = m^*(x)$.

Proof. Proof. (a) From Lemma 7.5(a, b) and $g(z, z) > 0$ yields

$$\Delta_2 F(p, z) = -f(z, z) \tilde{\mu}_p^{z^-} \{z\} = -m^*(z) g(z, z) \mu_p^{z^-} \{z\} = m^*(z) \Delta_2 G(p, z).$$

(b) Using in turn (19), part (a) and (20), gives

$$\Delta_2 f(x, z) = \beta \Delta_{a=0}^{a=1} \int \Delta_2 F(y, z) Q^a(x, dy) = m^*(z) \beta \Delta_{a=0}^{a=1} \int \Delta_2 G(y, z) Q^a(x, dy) = m^*(z) \Delta_2 g(x, z).$$

(c) Taking $z = x$ in part (b) and noting that $m^*(x) = f(x, x)/g(x, x)$, we obtain

$$f(x, x^-) - f(x, x) = m^*(x) \{g(x, x^-) - g(x, x)\} = m^*(x) g(x, x^-) - f(x, x).$$

whence $f(x, x^-) = m^*(x) g(x, x^-)$ and, by Lemma 7.6, $m^*(x) = m(x, x^-)$. \square

Lemma 7.8(a) yields the following alternate characterization of the MP index, under (PCLI1), in terms of the jumps $\Delta_2 F(x, x) = F(x, x) - F(x, x^-)$ and $\Delta_2 G(x, x) = G(x, x) - G(x, x^-)$.

Corollary 7.1. *Under (PCLI1), $m^*(x) = \Delta_2 F(x, x)/\Delta_2 G(x, x)$.*

We further immediately obtain the following result relating the continuity points of functions $F(x, \cdot)$ and $G(x, \cdot)$, and of $f(x, \cdot)$ and $g(x, \cdot)$.

Corollary 7.2. *Under (PCLI1),*

- (a) *if $G(p, \cdot)$ is continuous at z , then so is $F(p, \cdot)$;*
- (b) *if $g(x, \cdot)$ is continuous at z , then so is $f(x, \cdot)$.*

7.4 Bounded variation and MP index as Radon–Nikodým derivative.

We next show that the performance metrics of a PCL-indexable project, as functions of the threshold, belong to the linear space $\mathbb{V}(\mathbb{R})$ of functions of *bounded variation* on \mathbb{R} (see Carter and van Brunt [8, Sec. 2.7]). We further build on that result to characterize the MP index as a Radon–Nikodým derivative.

We start with the resource metrics, for which condition (PCLI1) suffices.

Lemma 7.9. *Under (PCLI1), $G(p, \cdot), g(x, \cdot) \in \mathbb{V}(\mathbb{R})$.*

Proof. Proof. $G(p, \cdot) \in \mathbb{V}(\mathbb{R})$ follows from $G(p, \cdot)$ being bounded (see (7)) and nonincreasing on \mathbb{R} (see Lemma 7.3(a)). Since $g(x, z) = \Delta_{a=0}^{a=1} \{c(x, a) + \beta G(Q^a(x, \cdot), z)\}$, $g(x, \cdot)$ is a difference of bounded nonincreasing functions, whence (see Carter and van Brunt [8, Th. 2.7.2]) $g(x, \cdot) \in \mathbb{V}(\mathbb{R})$. \square

Remark 7.1. Since the function $G(p, \cdot)$ is bounded, càdlàg and nonincreasing, by Carathéodory's extension theorem it induces a unique *finite LS measure* $\nu_{G(p, \cdot)}$ on $\mathcal{B}(\mathbb{R})$ satisfying

$$\nu_{G(p, \cdot)}(z_1, z_2] = G(p, z_1) - G(p, z_2), \quad -\infty < z_1 \leq z_2 < \infty.$$

In light of Remark 7.1, we have by standard results that the LS integral

$$\int |m^*|(z) G(p, dz) = - \int |m^*| d\nu_{G(p, \cdot)}$$

is well defined. The following result shows that such an integral is finite, even when \mathbf{X} is unbounded.

Lemma 7.10. *Under PCL-indexability, m^* is $\nu_{G(p,\cdot)}$ -integrable.*

Proof. Proof. In the case $m^* \geq 0$ over \mathbb{R} , we have

$$\begin{aligned} \int |m^*| d\nu_{G(p,\cdot)} &= \int m^* \nu_{G(p,\cdot)} = \lim_{n \rightarrow \infty} \int_{(-n,n]} m^* \nu_{G(p,\cdot)} \\ &= \lim_{n \rightarrow \infty} \{F(p, -n) - F(p, n)\} = F(p, -\infty) - F(p, \infty) < \infty, \end{aligned}$$

where we have used the monotone convergence theorem, condition (PCLI3) and Lemma 7.2(b, c). A similar argument yields the result in the case $m^* \leq 0$ over \mathbb{R} .

Otherwise, (PCLI2) implies that there exists a state b with $m^*(b) = 0$ such that $|m^*| = -m^*$ on $(-\infty, b]$ and $|m^*| = m^*$ on (b, ∞) . Arguing along the same lines as above yields

$$\begin{aligned} \int |m^*| \nu_{G(p,\cdot)} &= \int_{(-\infty,b]} -m^* d\nu_{G(p,\cdot)} + \int_{(b,\infty)} m^* \nu_{G(p,\cdot)} \\ &= \lim_{n \rightarrow \infty} \int_{(-n,b]} -m^* d\nu_{G(p,\cdot)} + \lim_{n \rightarrow \infty} \int_{(b,n]} m^* \nu_{G(p,\cdot)} \\ &= \lim_{n \rightarrow \infty} \{F(p, b) - F(p, -n)\} + \lim_{n \rightarrow \infty} \{F(p, b) - F(p, n)\} \\ &= 2F(p, b) - F(p, -\infty) - F(p, \infty) < \infty. \end{aligned}$$

□

The following result shows that, under PCL-indexability, the identity (28) —and, in particular condition (PCLI3) in Definition 2.4— extends to infinite intervals. Recall from §2.1 that the notation $(z_1, z_2]$ stands for the real interval $(z_1, z_2] \triangleq \{z \in \mathbb{R} : z_1 < z \leq z_2\}$ so, e.g., $(0, \infty] = (0, \infty)$.

Lemma 7.11. *Under PCL-indexability,*

$$F(p, z_2) - F(p, z_1) = \int_{(z_1, z_2]} m^*(z) G(p, dz), \quad -\infty \leq z_1 < z_2 \leq \infty. \quad (61)$$

Proof. Proof. The result follows from (28), Lemmas 7.2(b) and 7.10, and dominated convergence. □

We next present the analog of Lemma 7.9 for reward measures.

Lemma 7.12. *Under PCL-indexability, $F(p, \cdot), f(x, \cdot) \in \mathbb{V}(\mathbb{R})$.*

Proof. Proof. Since $\nu_{G(p,\cdot)}$ (see Remark 7.1) is a finite measure on $\mathcal{B}(\mathbb{R})$, Lemma 7.10 ensures that the set function $\tilde{\nu}_p(S) \triangleq \int_S m^* d\nu_{G(p,\cdot)}$ is a finite signed measure on $\mathcal{B}(\mathbb{R})$, and hence admits a Jordan decomposition $\tilde{\nu}_p = \tilde{\nu}_p^+ - \tilde{\nu}_p^-$ with $\tilde{\nu}_p^+$ and $\tilde{\nu}_p^-$ being finite measures. On the other hand, (PCLI3) and Lemma 7.2(b) yield that, for $z \in \mathbb{R}$, $\tilde{\nu}_p(-\infty, z] = F(p, -\infty) - F(p, z)$, whence $F(p, z) = F(p, -\infty) + \tilde{\nu}_p^+(-\infty, z] - \tilde{\nu}_p^-(-\infty, z]$. Hence, $F(p, \cdot) \in \mathbb{V}(\mathbb{R})$, being the difference of two bounded nondecreasing functions on \mathbb{R} . Further, it follows that the function mapping z to $F(\kappa^a(p, \cdot), z)$ is in $\mathbb{V}(\mathbb{R})$. Hence, the identity $f(x, z) = \Delta_{a=0}^{\frac{1}{2}} \{r(x, a) + \beta F(\kappa^a(x, \cdot), z)\}$ represents $f(x, \cdot)$ as the difference of two functions in $\mathbb{V}(\mathbb{R})$, whence $f(x, \cdot) \in \mathbb{V}(\mathbb{R})$. □

Remark 7.2. In light of Lemma 7.12 and $F(p, \cdot)$ being bounded càdlàg (see (8) and Lemma 7.2), under PCL-indexability the Carathéodory extension theorem for signed measures (cf. Doob [10, Sec. X.6]) yields that there exists a unique *finite LS signed measure* $\nu_{F(p, \cdot)}$ on $\mathcal{B}(\mathbb{R})$ satisfying

$$\nu_{F(p, \cdot)}(z_1, z_2] = F(p, z_1) - F(p, z_2), \quad -\infty < z_1 \leq z_2 < \infty.$$

The following result shows that the MP index of a PCL-indexable project is a Radon–Nikodým derivative of the signed measure $\nu_{F(p, \cdot)}$ with respect to the measure $\nu_{G(p, \cdot)}$. The notation \ll below means, as usual, “is absolutely continuous with respect to.”

Proposition 7.1. *Under PCL-indexability, $\nu_{F(p, \cdot)} \ll \nu_{G(p, \cdot)}$, with m^* being a Radon–Nikodým derivative of $\nu_{F(p, \cdot)}$ with respect to $\nu_{G(p, \cdot)}$, viz.,*

$$\nu_{F(p, \cdot)}(S) = \int_S m^* d\nu_{G(p, \cdot)}, \quad S \in \mathcal{B}(\mathbb{R}). \quad (62)$$

Proof. Proof. The result that $\nu_{F(p, \cdot)}$ satisfies (62) follows from (27) and Remark 7.2, whereas $\nu_{F(p, \cdot)} \ll \nu_{G(p, \cdot)}$ follows directly from (62). \square

By standard results on differentiation of LS measures (cf. Doob [10, §X.4]), the above ensures that m^* is a.s.- $\nu_{G(p, \cdot)}$ the *derivative of the signed measure $\nu_{F(p, \cdot)}$ with respect to the measure $\nu_{G(p, \cdot)}$* , denoted by $d\nu_{F(p, \cdot)}/\nu_{G(p, \cdot)}$, or of $F(p, \cdot)$ with respect to $G(p, \cdot)$, denoted by $dF(p, \cdot)/dG(p, \cdot)$. The next result gives a tighter characterization of m^* when p has full support.

Proposition 7.2. *Let PCL-indexability hold. If p has full support \mathbb{X} then*

$$m^*(z_0) = \frac{dF(p, \cdot)}{dG(p, \cdot)}(z_0) = \lim_{z \nearrow z_0} \frac{F(p, z) - F(p, z_0)}{G(p, z) - G(p, z_0)} = \lim_{z \nearrow z_0} \frac{F(p, z) - F(p, z_0^-)}{G(p, z) - G(p, z_0^-)}, \quad z_0 \in (\ell, u). \quad (63)$$

Proof. Proof. We focus on the second identity in (63). Consider first the left limit. By Lemma 7.3(b), $\nu_{G(p, \cdot)}(z, z_0] = G(p, z) - G(p, z_0) > 0$ for $z < z_0$. Let $\varepsilon > 0$. By (PCLI2), there exists $\delta > 0$ such that $m^*(b) \in (m^*(z_0) - \varepsilon, m^*(z_0)]$ for $b \in (z_0 - \delta, z_0]$. Hence, for any $z \in (z_0 - \delta, z_0]$ we have

$$\int_{(z, z_0]} (m^*(z_0) - \varepsilon) \nu_{G(p, \cdot)}(db) < \int_{(z, z_0]} m^*(b) \nu_{G(p, \cdot)}(db) \leq \int_{(z, z_0]} m^*(z_0) \nu_{G(p, \cdot)}(db),$$

which is readily reformulated as

$$m^*(z_0) - \varepsilon < \frac{\nu_{F(p, \cdot)}(z, z_0]}{\nu_{G(p, \cdot)}(z, z_0]} \leq m^*(z_0).$$

Therefore,

$$\lim_{z \nearrow z_0} \frac{F(p, z) - F(p, z_0)}{G(p, z) - G(p, z_0)} = \lim_{z \nearrow z_0} \frac{\nu_{F(p, \cdot)}(z, z_0]}{\nu_{G(p, \cdot)}(z, z_0]} = m^*(z_0).$$

Consider now the right limit. By Lemma 7.3(b), we have $\nu_{G(p, \cdot)}(z_0, z] = G(p, z_0) - G(p, z) > 0$ for $z > z_0$. Let $\varepsilon > 0$. By (PCLI2), there exists $\delta > 0$ such that $m^*(b) \in [m^*(z_0), m^*(z_0) + \varepsilon]$ for $b \in [z_0, z_0 + \delta)$. Hence, for any $z \in (z_0, z_0 + \delta)$ we have

$$\int_{(z_0, z]} m^*(z_0) \nu_{G(p, \cdot)}(db) \leq \int_{(z_0, z]} m^*(b) \nu_{G(p, \cdot)}(db) < \int_{(z_0, z]} (m^*(z_0) + \varepsilon) \nu_{G(p, \cdot)}(db),$$

i.e.,

$$m^*(z_0) \leq \frac{\nu_{F(p,\cdot)}(z_0, z]}{\nu_{G(p,\cdot)}(z_0, z]} < m^*(z_0) + \varepsilon.$$

Therefore,

$$\lim_{z \searrow z_0} \frac{F(p, z) - F(p, z_0)}{G(p, z) - G(p, z_0)} = \lim_{z \searrow z_0} \frac{\nu_{F(p,\cdot)}(z_0, z]}{\nu_{G(p,\cdot)}(z_0, z]} = m^*(z_0).$$

As for the third identity in (63), it follows along the same lines as the second identity, using the intervals (z, z_0) for $z < z_0$ and $[z_0, z]$ for $z > z_0$. \square

7.5 Relations between marginal performance metrics.

This section presents relations between marginal performance metrics as functions of the threshold, which we will use in the sequel. The following result gives counterparts of Lemmas 7.10 and 7.11 for the marginal metrics $f(x, \cdot)$ and $g(x, \cdot)$. Note that the LS integrals below are well defined, since $g(x, \cdot)$ is a bounded function of bounded variation on \mathbb{R} and m^* is continuous. The result refers to the finite signed measure $\nu_{g(x,\cdot)}$ determined by $g(x, \cdot)$ through $\nu_{g(x,\cdot)}(z_1, z_2] = g(x, z_1) - g(x, z_2)$ for finite $z_1 \leq z_2$, which satisfies .

Lemma 7.13. *Under PCL-indexability,*

- (a) m^* is $\nu_{g(x,\cdot)}$ -integrable, i.e., the integral $\int |m^*| d|\nu_{g(x,\cdot)}|$ is finite;
- (b) $f(x, z_2) - f(x, z_1) = \int_{(z_1, z_2]} m^*(z) g(x, dz)$, for $-\infty \leq z_1 < z_2 \leq \infty$.

Proof. Proof. (a) This part follows from Lemma 7.10 using that, by (24), $\nu_{g(x,\cdot)} = \beta \Delta_{a=0}^{a=1} \nu_{G(\kappa^a(x,\cdot), \cdot)}$.

(b) Using part (a), Lemma 7.11 and (24), we obtain

$$\begin{aligned} \int_{(z_1, z_2]} m^*(z) g(x, dz) &= \beta \Delta_{a=0}^{a=1} \int_{(z_1, z_2]} m^*(z) G(\kappa^a(x, \cdot), dz) \\ &= \beta \Delta_{a=0}^{a=1} \{F(\kappa^a(x, \cdot), z_2) - F(\kappa^a(x, \cdot), z_1)\} = f(x, z_2) - f(x, z_1). \end{aligned}$$

\square

The next result shows that m^* satisfies certain linear Volterra–Stieltjes integral equations.

Lemma 7.14. *Under PCL-indexability, for any threshold $z \in \overline{\mathbb{R}}$,*

- (a) if $x > z$, $m(x, z) - m^*(z) = \int_{(z, x)} \frac{g(x, b)}{g(x, z)} m^*(db)$; in particular,

$$m(x, \ell^-) - m^*(\ell) = \int_{[\ell, x)} \frac{g(x, b)}{g(x, \ell^-)} m^*(db); \quad (64)$$

- (b) if $x \leq z$, $m(x, z) - m^*(z) = -\int_{[x, z]} \frac{g(x, b)}{g(x, z)} m^*(db)$.

Proof. Proof. (a) From Lemmas 7.2, 7.13, and the dominated convergence theorem, we have

$$f(x, x^-) - f(x, z) = \lim_{y \nearrow x} f(x, y) - f(x, z) = \lim_{y \nearrow x} \int_{(z, y]} m^*(b) g(x, db) = \int_{(z, x)} m^*(b) g(x, db).$$

Now, using the latter identity, integration by parts, and Lemma 7.8(c) gives

$$\begin{aligned} f(x, z) &= f(x, x^-) - \int_{(z, x)} m^*(b) g(x, db) \\ &= f(x, x^-) - \{m^*(x)g(x, x^-) - m^*(z)g(x, z)\} + \int_{(z, x)} g(x, b) m^*(db) \\ &= m^*(z)g(x, z) + \int_{(z, x)} g(x, b) m^*(db). \end{aligned}$$

The result now follows by dividing each term by $g(x, z)$.

(b) Consider first the finite threshold case $z < \infty$. Arguing as in part (a), we obtain

$$\begin{aligned} f(x, z) &= f(x, x^-) + \int_{[x, z]} m^*(b) g(x, db) \\ &= f(x, x^-) + \{m^*(z)g(x, z) - m^*(x)g(x, x^-)\} - \int_{[x, z]} g(x, b) m^*(db) \\ &= m^*(z)g(x, z) - \int_{[x, z]} g(x, b) m^*(db). \end{aligned} \tag{65}$$

In the infinite threshold case, the result follows from the limiting argument

$$\begin{aligned} f(x, \infty) - m^*(\infty)g(x, \infty) &= \lim_{z \rightarrow \infty} \{f(x, z) - m^*(z)g(x, z)\} \\ &= - \lim_{z \rightarrow \infty} \int_{[x, z]} g(x, b) m^*(db) = - \int_{[x, \infty)} g(x, b) m^*(db), \end{aligned}$$

which applies whether $m^*(\infty) < \infty$ or $m^*(\infty) = \infty$, where we have used in turn Lemma 7.2(c), (PCL1), (65) and (PCL2). This completes the proof. \square

The next result (cf. Lemma 3.3), which follows immediately from Lemma 7.14, shows that the MP metric $m(x, z)$ yields lower and upper bounds on the MP index $m^*(z)$. Recall that the “sup” and the “inf” over the empty set are taken to be $-\infty$ and ∞ , respectively, as in (31).

Corollary 7.3. *Under PCL-indexability,*

$$\sup_{x \in [\ell, z]} m(x, z) \leq m^*(z) \leq \inf_{x \in (z, u]} m(x, z), \quad z \in \overline{\mathbb{R}}. \tag{66}$$

8 Dual LP and strong duality.

In this section we solve the *primal* LP relaxation $R_\lambda(x)$ in (57) under PCL-indexability, via LP duality (cf. Anderson and Nash [1, pp. 38–39]). We will postulate, justify and solve in closed form a dual LP to $R_\lambda(x)$, establishing strong duality.

The variables of the dual to $R_\lambda(x)$ are a function $\psi: \mathbf{X} \rightarrow \mathbb{R}$ and two scalars, η and ξ . See (57). We will use ψ 's from the *convex cone* $\mathbb{CV}^{\text{ND}}(\mathbf{X})$ of nondecreasing functions in $\mathbb{CV}(\mathbf{X}) \triangleq \mathbb{C}(\mathbf{X}) \cap \mathbb{V}(\mathbf{X})$, the

vector space of real-valued continuous functions with bounded variation on \mathbf{X} (see §7.4). Note that $\mathbb{CV}^{\text{ND}}(\mathbf{X})$ consists of the continuous, nondecreasing and bounded functions on \mathbf{X} .

We posit the following LP as a dual to $R_\lambda(x)$ (where \mathbb{R}^+ denotes the nonnegative reals):

$$\begin{aligned}
D_\lambda(x): \quad & \text{maximize } \int_{[\ell, u)} G(x, b) \psi(db) + G(x, \ell^-) \eta + G(x, u) \xi \\
& \text{subject to: } \psi \in \mathbb{CV}^{\text{ND}}(\mathbf{X}), \eta \in \mathbb{R}, \xi \in \mathbb{R}^+ \\
\tilde{\mu}^0(dy): \quad & \int_{[\ell, y)} g(y, b) \psi(db) + g(y, \ell^-) \eta \leq f(y, \ell^-), \quad y \in \mathbf{X} \\
\gamma: \quad & \int_{[\ell, u)} \psi(db) + \eta + \xi = \lambda.
\end{aligned} \tag{67}$$

Note that we have indicated beside each constraint the corresponding primal variable, and that using integrators $\psi \in \mathbb{CV}^{\text{ND}}(\mathbf{X})$ ensures that the LS integrals in (67) exist and are finite, since the integrands have bounded variation on \mathbb{R} . See §7.4. Note further that, in the case $\ell > -\infty$, the constraint in (67) corresponding to $y = \ell$ reduces to $g(\ell, \ell^-) \eta \leq f(\ell, \ell^-)$, since $[\ell, \ell) = \emptyset$.

Denote by $\text{obj}(\psi, \eta, \xi; D(x))$ the objective value in (67) of a $D_\lambda(x)$ -feasible solution (ψ, η, ξ) . Recall that $\text{obj}(\tilde{\mu}^0, \gamma; R_\lambda)$ is the primal objective value in (57) of an $R_\lambda(x)$ -feasible solution $(\tilde{\mu}^0, \gamma)$. Write as $\Delta_{x, \lambda}(\psi, \eta, \xi; \tilde{\mu}^0, \gamma) \triangleq \text{obj}(\tilde{\mu}^0, \gamma; R_\lambda) - \text{obj}(\psi, \eta, \xi; D(x))$ the corresponding *gap*. The following result establishes that the primal-dual pair of LPs $R_\lambda(x)$ and $D_\lambda(x)$ satisfy *weak duality*.

Lemma 8.1 (Weak duality). *Let (PCLI1) hold. Then $\Delta_{x, \lambda}(\psi, \eta, \xi; \tilde{\mu}^0, \gamma) \geq 0$.*

Proof. Proof. We can decompose $\Delta_{x, \lambda}(\psi, \eta, \xi; \tilde{\mu}^0, \gamma)$ through algebraic manipulations as

$$\begin{aligned}
\Delta_{x, \lambda}(\psi, \eta, \xi; \tilde{\mu}^0, \gamma) &= \int \left\{ f(y, \ell^-) - \int_{[\ell, y)} g(y, b) \psi(db) - g(y, \ell^-) \eta \right\} \tilde{\mu}^0(dy) \\
&\quad + \int_{[\ell, u)} \left\{ \int_{(b, u]} g(y, b) \tilde{\mu}^0(dy) + \gamma - G(x, b) \right\} \psi(db) + \{\gamma - G(x, u)\} \xi,
\end{aligned} \tag{68}$$

where, to cancel terms, we have used that $\int g(y, \ell^-) \tilde{\mu}^0(dy) + \gamma = G(x, \ell^-)$ and $\int_{[\ell, u)} \psi(db) + \eta + \xi = \lambda$ (from the constraints of $R_\lambda(x)$ and $D_\lambda(x)$), and Tonelli's theorem, applying that, since $\psi \in \mathbb{CV}^{\text{ND}}(\mathbf{X})$, $[\ell, \ell) = \emptyset$, and $g(x, b) > 0$ by (PCLI1),

$$\int \int_{[\ell, y)} g(y, b) \psi(db) \tilde{\mu}^0(dy) = \int_{(\ell, u]} \int_{[\ell, y)} g(y, b) \psi(db) \tilde{\mu}^0(dy) = \int_{[\ell, u)} \int_{(b, u]} g(y, b) \tilde{\mu}^0(dy) \psi(db).$$

The nonnegativity of $\Delta_{x, \lambda}(\psi, \eta, \xi; \tilde{\mu}^0, \gamma)$ follows because each integrand in (68) is nonnegative, $\tilde{\mu}^0$ is a measure, ψ is nondecreasing, $\gamma \geq G(x, u)$ and $\xi \geq 0$. \square

From (68) we obtain the following *complementary slackness* (CS) conditions for a primal-dual feasible solution pair $(\tilde{\mu}^0, \gamma)$ and (ψ, η, ξ) , which characterize their optimality:

$$\int \left\{ f(y, \ell^-) - \int_{[\ell, y)} g(y, b) \psi(db) - g(y, \ell^-) \eta \right\} \tilde{\mu}^0(dy) = 0 \tag{69}$$

$$\int_{[\ell, u)} \left\{ \int_{(b, u]} g(y, b) \tilde{\mu}^0(dy) + \gamma - G(x, b) \right\} \psi(db) = 0 \tag{70}$$

$$\{\gamma - G(x, u)\} \xi = 0. \tag{71}$$

Lemma 8.2 (CS optimality conditions). *Under (PCLI1), a primal-dual feasible solution pair $(\tilde{\mu}^0, \gamma)$ and (ψ, η, ξ) is optimal for $R_\lambda(x)$ and $D_\lambda(x)$, respectively, iff it satisfies (69–71).*

Proof. Proof. The result follows immediately from Lemma 8.1, (68) and feasibility. \square

We next set out to show that the LPs $R_\lambda(x)$ and $D_\lambda(x)$ satisfy *strong duality*, i.e., they have the same optimal value, and this is attained in each problem. We will construct a feasible dual solution $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ for each resource price $\lambda \in \mathbb{R}$. Then, we will show that there exists a threshold z such that, for any initial state x , the primal feasible solution $(\tilde{\mu}_x^{z,0}, G(x, z))$ satisfies CS with $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$. Lemma 8.2 will then ensure the optimality of such solutions.

For a resource price $\lambda \in \mathbb{R}$, let

$$\psi_\lambda \triangleq \min\{m^*, \lambda\}, \quad \eta_\lambda \triangleq \psi_\lambda(\ell), \quad \xi_\lambda \triangleq \lambda - \psi_\lambda(u) = \{\lambda - m^*(u)\}^+, \quad (72)$$

where $y^+ \triangleq \max\{y, 0\}$. Note that $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ is well defined under PCL-indexability conditions (PCLI1, PCLI2) in Definition 2.4, even when $\ell = -\infty$ or $u = \infty$.

Lemma 8.3. *Under PCL-indexability, $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ is $D_\lambda(x)$ -feasible.*

Proof. Proof. From (PCLI1, PCLI2) and (72) we obtain $\psi_\lambda \in \mathbb{C}\mathbb{V}^{\text{ND}}(\mathbf{X})$ and $\xi_\lambda \geq 0$. Further, $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ satisfies the second constraint of $D_\lambda(x)$, since

$$\int_{[\ell, u]} \psi_\lambda(db) + \eta_\lambda + \xi_\lambda = \psi_\lambda(u) - \psi_\lambda(\ell) + \psi_\lambda(\ell) + \lambda - \psi_\lambda(u) = \lambda. \quad (73)$$

To verify that $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ satisfies the first constraint of $D_\lambda(x)$, we distinguish three cases.

Case I: $\lambda < m^*(\ell)$. Then, $\psi_\lambda \equiv \lambda$, $\eta_\lambda = \lambda$, $\xi_\lambda = 0$ and, for any state y ,

$$\begin{aligned} f(y, \ell^-) - \int_{[\ell, y]} g(y, b) \psi_\lambda(db) - g(y, \ell^-) \eta_\lambda &= f(y, \ell^-) - \lambda g(y, \ell^-) \\ &= \int_{[\ell, y]} g(y, b) m^*(db) + \{m^*(\ell) - \lambda\} g(y, \ell^-) > 0, \end{aligned} \quad (74)$$

where we have used (64), (PCLI1, PCLI2) and $\lambda < m^*(\ell)$.

Case II: $\lambda > m^*(u)$. Then, $m^*(u)$ is finite, $\psi_\lambda = m^*$, $\eta_\lambda = m^*(\ell)$, and $\xi_\lambda = \lambda - m^*(u)$. Hence, for any state y , (64) yields

$$\int_{[\ell, y]} g(y, b) \psi_\lambda(db) + g(y, \ell^-) \eta_\lambda = \int_{[\ell, y]} g(y, b) m^*(db) + g(y, \ell^-) m^*(\ell) = f(y, \ell^-). \quad (75)$$

Case III: $m^*(\ell) \leq \lambda \leq m^*(u)$. Then, $\eta_\lambda = m^*(\ell)$ and $\xi_\lambda = 0$. Further, letting

$$\bar{y}_\lambda \triangleq \sup\{y \in \mathbf{X} : m^*(y) \leq \lambda\}, \quad (76)$$

by (PCLI2) we have $\bar{y}_\lambda \in \mathbf{X}$, with $\psi_\lambda(y) = m^*(y)$ for states $y \leq \bar{y}_\lambda$ and $\psi_\lambda(y) \equiv \lambda$ for states $y \geq \bar{y}_\lambda$. Hence, for $y \leq \bar{y}_\lambda$ we can use (64) to obtain (75).

Further, for any state $y > \bar{y}_\lambda$ we can use (PCLI1, PCLI2) and (64) to obtain

$$\begin{aligned} \int_{[\ell, y]} g(y, b) \psi_\lambda(db) + g(y, \ell^-) \eta_\lambda &= \int_{[\ell, \bar{y}_\lambda]} g(y, b) m^*(db) + m^*(\ell) g(y, \ell^-) \\ &< \int_{[\ell, y]} g(y, b) m^*(db) + m^*(\ell) g(y, \ell^-) = f(y, \ell^-), \end{aligned}$$

where the “ $<$ ” follows from $\int_{(\bar{y}_\lambda, y)} g(y, b) m^*(db) > 0$, since $m^*(y) > \lambda = m^*(\bar{y}_\lambda)$. \square

Note (cf. Lemmas 6.1 and 6.2) that, for each threshold z and state x , $(\tilde{\mu}_x^{z,0}, G(x, z))$ is $R_\lambda(x)$ -feasible. For $\lambda \in \mathbb{R}$, let \hat{Z}_λ^* be the set of thresholds z for which $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ satisfies CS conditions (69–71) with $(\tilde{\mu}_x^{z,0}, G(x, z))$ for every x . The next result characterizes \hat{Z}_λ^* through the index m^* .

Lemma 8.4. *Under PCL-indexability, \hat{Z}_λ^* is the nonempty interval of $\overline{\mathbb{R}}$ given by*

$$\hat{Z}_\lambda^* = \begin{cases} \{z \in \overline{\mathbb{R}}: -\infty \leq z < \ell\} & \text{if } \lambda < m^*(\ell); \\ \{z \in \overline{\mathbb{R}}: u \leq z \leq \infty\} & \text{if } \lambda > m^*(u); \\ \{z \in \overline{\mathbb{R}}: m^*(z) = \lambda\} & \text{if } m^*(\ell) \leq \lambda \leq m^*(u). \end{cases}$$

Proof. Proof. We distinguish three cases, as in the proof of Lemma 8.3.

Case I: $\lambda < m^*(\ell)$. Then, $\psi_\lambda \equiv \lambda$, $\eta_\lambda = \lambda$ and $\xi_\lambda = 0$, whence (69) reduces to

$$\int \{f(y, \ell^-) - \lambda g(y, \ell^-)\} \tilde{\mu}_x^{z,0}(dy) = 0. \quad (77)$$

Since, by (74), $f(y, \ell^-) - \lambda g(y, \ell^-) > 0$ for each y , (77) holds for every x iff the z -policy takes the active action in every state, i.e., iff $z < \ell$. As for (70), it holds for every z , because $\psi_\lambda \equiv \lambda$. Condition (71) also holds for every z . Hence, $\hat{Z}_\lambda^* = \{z \in \overline{\mathbb{R}}: -\infty \leq z < \ell\}$.

Case II: $\lambda > m^*(u)$. Then, $\psi_\lambda = m^*$, $\eta_\lambda = m^*(\ell)$ and $\xi_\lambda = \lambda - m^*(u)$. Regarding (69), it holds for any z , because the integrand in (69) vanishes everywhere. See (75). As for (70), it reduces to

$$\int_{[\ell, u]} \left\{ G(x, z) + \int_{(b, u]} g(y, b) \tilde{\mu}_x^{z,0}(dy) - G(x, b) \right\} m^*(db) = 0. \quad (78)$$

If $z \geq u$, the z -policy always rests the project. Hence, by Lemma 6.1 and PCL(c.2) in Definition 6.1, the integrand in the left hand side of (78) vanishes for every x , whence (78) holds.

Concerning CS condition (71), we can formulate it as

$$\{G(x, z) - G(x, u)\} \xi_\lambda = \{G(x, z) - G(x, u)\} \{\lambda - m^*(u)\} = 0,$$

whence it holds iff $G(x, z) = G(x, u)$. Now, it follows from Lemma 6.1 and PCL(b.2) in Definition 6.1 that $G(x, z) = G(x, u)$ for every x iff $z \geq u$. Hence, $\hat{Z}_\lambda^* = \{z \in \overline{\mathbb{R}}: u \leq z \leq \infty\}$.

Case III: $m^*(\ell) \leq \lambda \leq m^*(u)$. Then, $\eta_\lambda = m^*(\ell)$, $\xi_\lambda = 0$, $\psi_\lambda(y) = m^*(y)$ for states $y \leq \bar{y}_\lambda$, and $\psi_\lambda(y) \equiv \lambda$ for $y \geq \bar{y}_\lambda$, with \bar{y}_λ as in (76). We start by showing that $(\tilde{\mu}_x^{z,0}, G(x, z))$ satisfies (69) with $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ for any x iff $m^*(z) \leq \lambda$. So assume (69) holds for any x and, to argue by contradiction, suppose $m^*(z) > \lambda$. It would then follow from (PCLI2) that $\bar{y}_\lambda < \min\{u, z\}$ and, for any $x \in (\bar{y}_\lambda, z]$,

$$\begin{aligned} 0 &= \int \left\{ f(y, \ell^-) - \int_{[\ell, y]} g(y, b) \psi_\lambda(db) - g(y, \ell^-) \eta_\lambda \right\} \tilde{\mu}_x^{z,0}(dy) \\ &= \int_{(\bar{y}_\lambda, u]} \left\{ f(y, \ell^-) - \int_{[0, \bar{y}_\lambda]} g(y, b) m^*(db) - m^*(\ell) g(y, \ell^-) \right\} \tilde{\mu}_x^{z,0}(dy) \\ &= \int_{(\bar{y}_\lambda, u]} \int_{[\bar{y}_\lambda, y]} g(y, b) m^*(db) \tilde{\mu}_x^{z,0}(dy) \geq \int_{[\bar{y}_\lambda, x]} g(x, b) m^*(db) > 0, \end{aligned}$$

a contradiction, where we have used (64), (PCLI1, PCLI2), $\tilde{\mu}_x^{z,0}\{x\} \geq 1$, and $m^*(\bar{y}_\lambda) = \lambda < m^*(x)$. Hence, it must be $m^*(z) \leq \lambda$.

Conversely, suppose $m^*(z) \leq \lambda$. If $z < \ell$, then $\tilde{\mu}_x^{z,0}(\mathbf{X}) = 0$, whence (69) holds. If $z \geq \ell$, then, letting $\underline{z} \triangleq \min\{z, u\}$, we have

$$\begin{aligned} & \int \left\{ f(y, \ell^-) - \int_{[\ell, y)} g(y, b) \psi_\lambda(db) - g(y, \ell^-) \eta_\lambda \right\} \tilde{\mu}_x^{z,0}(dy) \\ &= \int_{[\ell, \underline{z}]} \left\{ f(y, \ell^-) - \int_{[\ell, y)} g(y, b) m^*(db) - g(y, \ell^-) m^*(\ell) \right\} \tilde{\mu}_x^{z,0}(dy) = 0, \end{aligned}$$

where we have used that $\tilde{\mu}_x^{z,0}(\underline{z}, u] = 0$ and $f(\ell, \ell^-) = g(\ell, \ell^-) m^*(\ell)$ if ℓ is finite (by Lemma 7.8(c)). Hence (69) holds for every x .

As for (70), we will show it holds for every x iff $m^*(z) \geq \lambda$. We can formulate (70) as

$$\begin{aligned} 0 &= \int_{[\ell, u]} \left\{ \int_{(b, u]} g(y, b) \tilde{\mu}_x^{z,0}(dy) + G(x, z) - G(x, b) \right\} \psi_\lambda(db) \\ &= \int_{[\ell, \bar{y}_\lambda]} \left\{ \int_{(b, u]} g(y, b) \tilde{\mu}_x^{z,0}(dy) + G(x, z) - G(x, b) \right\} m^*(db) \\ &= \int_{[\ell, \bar{y}_\lambda]} \int_{[\ell, b]} g(y, b) \tilde{\mu}_x^{z,1}(dy) m^*(db) = \int_{(z, \bar{y}_\lambda)} \int_{[y, \bar{y}_\lambda]} g(y, b) m^*(db) \tilde{\mu}_x^{z,1}(dy), \end{aligned} \tag{79}$$

where we have used (56), Tonelli's theorem and $\tilde{\mu}_x^{z,1}[\ell, z] = 0$. Suppose $m^*(z) \geq \lambda$. If $z \leq \ell$, then $\psi_\lambda \equiv \lambda$ and (79) follows. If $\ell < z \geq \bar{y}_\lambda$, then (79) follows. If $\ell < z < \bar{y}_\lambda$ then $m^*(y) = \lambda$ for $y \in (z, \bar{y}_\lambda)$, whence (79) also holds.

Conversely, assume (79) holds for every x , and suppose it were $m^*(z) < \lambda$. As before, we can assume $z > \ell$. Then, $z < \bar{y}_\lambda$ and, for any $x \in (z, \bar{y}_\lambda)$ with $m^*(x) < \lambda$,

$$0 = \int_{(z, \bar{y}_\lambda)} \int_{[y, \bar{y}_\lambda]} g(y, b) m^*(db) \tilde{\mu}_x^{z,1}(dy) \geq \int_{[x, \bar{y}_\lambda]} g(x, b) m^*(db) > 0,$$

a contradiction. Hence, it must be $m^*(z) \geq \lambda$.

As for CS condition (71), it holds for any z , because $\xi_\lambda = 0$. Therefore, $\hat{\mathbf{Z}}_\lambda^* = \{z \in \overline{\mathbb{R}} : m^*(z) = \lambda\}$ in this case. This completes the proof. \square

The following result establishes that, for any price λ and state x , the primal problem $R_\lambda(x)$ and its dual $D_\lambda(x)$ satisfy strong duality.

Lemma 8.5 (Strong duality). *Under PCL-indexability,*

- (a) $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ is $D_\lambda(x)$ -optimal;
- (b) for any threshold $z \in \hat{\mathbf{Z}}_\lambda^*$, $(\tilde{\mu}_x^{z,0}, G(x, z))$ is $R_\lambda(x)$ -optimal;
- (c) for any threshold $z \in \hat{\mathbf{Z}}_\lambda^*$, $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ and $(\tilde{\mu}_x^{z,0}, G(x, z))$ have the same objective value, i.e., $\text{obj}(\tilde{\mu}_x^{z,0}, G(x, z); R_\lambda) = \text{obj}(\psi_\lambda, \eta_\lambda, \xi_\lambda; D(x))$, and so $\text{val}(R_\lambda(x)) = \text{val}(D_\lambda(x))$.

Proof. Proof. The results follow from Lemmas 8.1–8.4, as $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ and $(\tilde{\mu}_x^{z,0}, G(x, z))$ are feasible solutions to $D_\lambda(x)$ and $R_\lambda(x)$ satisfying CS conditions (69–71). \square

9 Optimality of threshold policies.

This section establishes, under PCL-indexability, the optimality of threshold policies for each λ -price problem (see (4)) in the parametric collection $\{P_\lambda: \lambda \in \mathbb{R}\}$, and further characterizes (along with Lemma 8.4) in terms of the MP index m^* the optimal threshold set Z_λ^* for P_λ and the optimal price set Λ_z^* . Recall (cf. Lemma 3.3) that Z_λ^* is the set of thresholds z such that the z -policy is P_λ -optimal, whereas $\Lambda_z^* \triangleq \{\lambda: z \in Z_\lambda^*\}$. Let $\lambda \in \mathbb{R}$.

We start with a preliminary result, which shows that problem $R_\lambda(x)$ is an *exact relaxation* of problem $\tilde{L}_\lambda(x)$, in the sense that both have the same optimal value. The result goes further by identifying a common optimal solution to both problems, given in terms of a threshold. Let $\lambda \in \mathbb{R}$.

Lemma 9.1 (Exact relaxation). *Suppose PCL-indexability holds, and let $z \in \hat{Z}_\lambda^*$. Then, for each state x , $(\tilde{\mu}_x^{z,0}, G(x, z))$ is an optimal solution to both $R_\lambda(x)$ and $\tilde{L}_\lambda(x)$.*

Proof. Proof. By Lemmas 8.4 and 8.5(b) there exists a threshold $z \in \hat{Z}_\lambda^*$ such that $(\tilde{\mu}_x^{z,0}, G(x, z))$ is $R_\lambda(x)$ -optimal. Since $(\tilde{\mu}_x^{z,0}, G(x, z))$ is $\tilde{L}_\lambda(x)$ -feasible, and by Lemma 6.2 $R_\lambda(x)$ is a relaxation of $\tilde{L}_\lambda(x)$, it follows that $(\tilde{\mu}_x^{z,0}, G(x, z))$ is also $\tilde{L}_\lambda(x)$ -optimal. \square

The following result establishes the optimality of threshold policies for each λ -price problem, and further characterizes the P_λ -optimal threshold policies. Recall that $\hat{Z}_\lambda^* \neq \emptyset$ by Lemma 8.4.

Lemma 9.2 (Optimality of threshold policies). *Under PCL-indexability, $Z_\lambda^* = \hat{Z}_\lambda^*$.*

Proof. Proof. Let $z \in \hat{Z}_\lambda^*$. Lemma 9.1 ensures that, for each state x , $(\tilde{\mu}_x^{z,0}, G(x, z))$ is optimal for both problems $R_\lambda(x)$ and $\tilde{L}_\lambda(x)$, whence the z -policy is $P_\lambda(x)$ -optimal. Therefore, $z \in Z_\lambda^*$.

Conversely, let $z \in Z_\lambda^*$, so the z -policy is P_λ -optimal, being hence optimal for $P_\lambda(x)$ and $\tilde{P}_\lambda(x)$ for each initial state x . Hence, $(\tilde{\mu}_x^{z,0}, G(x, z))$ is $\tilde{L}_\lambda(x)$ -optimal and, by Lemma 9.1, is also $R_\lambda(x)$ -optimal. It follows from strong duality (Lemma 8.5) that $(\tilde{\mu}_x^{z,0}, G(x, z))$ must satisfy CS with the optimal dual solution $(\psi_\lambda, \eta_\lambda, \xi_\lambda)$ for every x , whence $z \in \hat{Z}_\lambda^*$. This completes the proof. \square

The next result characterizes the optimal price set for a given threshold z .

Lemma 9.3. *Under PCL-indexability,*

$$\Lambda_z^* = \begin{cases} (-\infty, m^*(\ell)], & \text{if } z < \ell; \\ \{m^*(z)\}, & \text{if } \ell \leq z < u; \\ [m^*(u), \infty), & \text{if } z \geq u \text{ and } m^*(u) < \infty; \\ \emptyset, & \text{if } z = u = \infty \text{ and } m^*(u) = \infty. \end{cases}$$

Proof. Proof. The result follows from Lemmas 8.4 and 9.2, and $\Lambda_z^* = \{\lambda \in \mathbb{R}: z \in Z_\lambda^*\}$. \square

We will need in the sequel the following immediate consequence of Lemma 9.3.

Corollary 9.1. *Under PCL-indexability, for any price λ and threshold $z \in Z_\lambda^*$,*

- (a) $m^*(z) \geq \lambda$ if $z < u$;
- (b) $m^*(z) \leq \lambda$ if $z \geq \ell$.

10 Further relations between MP metrics.

Recall that Corollary 7.3 shows that the MP measure $m(x, z)$ gives lower and upper bounds on the MP index $m^*(z)$. The following result establishes the tightness of such bounds, providing a dual characterization of $m^*(z)$.

Lemma 10.1. *Under PCL-indexability,*

- (a) if $\ell = -\infty$, $m^*(-\infty) = \inf_{x \in \mathbf{X}} m(x, -\infty)$;
- (b) for $z \in \mathbf{X} \setminus \{u\}$, $\max_{x \in [\ell, z]} m(x, z) = m^*(z) = \inf_{x \in (z, u]} m(x, z)$;
- (c) $m^*(u) = \sup_{x \in \mathbf{X}} m(x, u)$, with “max” instead of “sup” if $u < \infty$.

Proof. Proof. The result follows by matching the expressions for Λ_z^* in Lemmas 3.3(b) and 9.3. \square

Below, $\text{sgn}(y) \in \{-1, 0, 1\}$ denotes the sign of y .

Lemma 10.2. *Under PCL-indexability, for any state x and threshold z ,*

$$\text{sgn}(m(x, z) - m^*(z)) = \text{sgn}(m^*(x) - m^*(z)).$$

Proof. Proof. If $x \leq z$, we use Lemma 7.14(b) and (PCLI1, PCLI2) to obtain

$$\text{sgn}(m(x, z) - m^*(z)) = -\text{sgn} \int_{[x, z]} g(x, b) m^*(db) = \text{sgn}(m^*(x) - m^*(z)) \in \{-1, 0\}.$$

If $x > z$, we use Lemma 7.14(c) and (PCLI1, PCLI2) to obtain

$$\text{sgn}(m(x, z) - m^*(z)) = \text{sgn} \int_{(z, x)} g(x, b) m^*(db) = \text{sgn}(m^*(x) - m^*(z)) \in \{0, 1\}.$$

\square

We need a final preliminary result.

Lemma 10.3. *Under PCL-indexability, for any $x \in \mathbf{X}$, $\lambda \in \mathbb{R}$ and $z \in \mathbf{Z}_\lambda^*$,*

$$\text{sgn}(m(x, z) - \lambda) = \text{sgn}(m^*(x) - \lambda).$$

Proof. Proof. We distinguish three cases. In the case $\ell \leq z < u$, we have by Lemma 9.3 that $m^*(z) = \lambda$, whence the result follows from Lemma 10.2.

In the case $z < \ell$, we have (i) $m^*(x) \geq m^*(\ell) = m^*(z) \geq \lambda$, by (PCLI2) and Corollary 9.1(a); (ii) $m(x, z) \geq \lambda$, by Lemma 3.3(a); and (iii) $m(x, z) \geq m^*(z)$, by Lemma 10.1(a). Further, if $m^*(x) = \lambda$ then (i) yields $m^*(x) = m^*(z) = \lambda$, whence $m(x, z) - \lambda = m(x, z) - m^*(z) = 0$ by Lemma 10.2. Conversely, if $m(x, z) = \lambda$ then (i) and (iii) yield $m(x, z) = m^*(z) = \lambda$, whence $m^*(x) = m^*(z) = \lambda$ by Lemma 10.2. This shows that $\text{sgn}(m(x, z) - \lambda) = \text{sgn}(m^*(x) - \lambda) \in \{0, 1\}$.

In the case $z > u$, (i) $m^*(x) \leq m^*(u) = m^*(z) \leq \lambda$, by (PCLI2) and Lemma 9.1(b); (ii) $m(x, z) \leq \lambda$, by Lemma 3.3(a); and (iii) $m(x, z) \leq m^*(z)$, by Lemma 10.1(c). Arguing as in the previous case yields $m^*(x) = \lambda$ iff $m(x, z) = \lambda$, whence $\text{sgn}(m(x, z) - \lambda) = \text{sgn}(m^*(x) - \lambda) \in \{-1, 0\}$. \square

11 Proof of Theorem 2.1.

We are now ready to prove Theorem 2.1.

Proof. Proof.[Theorem 2.1]. Let $\lambda \in \mathbb{R}$. By Lemma 9.2, there is a P_λ -optimal threshold policy $z \in Z_\lambda^*$. From Lemma 3.2, (PCLI1) and Lemma 10.3, we obtain the following: for each state x ,

$$x \in S_\lambda^{*,1} \iff f(x, z) - \lambda g(x, z) \geq 0 \iff m(x, z) \geq \lambda \iff m^*(x) \geq \lambda$$

and

$$x \in S_\lambda^{*,0} \iff f(x, z) - \lambda g(x, z) \leq 0 \iff m(x, z) \leq \lambda \iff m^*(x) \leq \lambda,$$

which shows that the project is indexable with index m^* (see Definition 2.1). The result that the project is strongly threshold-indexable then follows from Lemma 3.1 and (PCLI2). \square

12 Examples.

12.1 Optimal stopping and Gittins index.

The optimal stopping model corresponds to the special case where passive transitions do not change the project state. To simplify the analysis we assume without loss of generality that both passive rewards and resource consumptions are zero, i.e., $r(x, 0) = c(x, 0) \equiv 0$. Then it is easily seen that the reward and resource usage performance metrics are determined by the following relations: $F(x, z) = G(x, z) = 0$ for $x \leq z$, and

$$F(x, z) = r(x, 1) + \beta \int_{(z, u]} F(y, z) Q^1(x, dy), \quad G(x, z) = c(x, 1) + \beta \int_{(z, u]} G(y, z) Q^1(x, dy), \quad x > z.$$

As for the marginal performance metrics, they are given by

$$f(x, z) = \begin{cases} r(x, 1) + \beta \int_{(z, u]} F(y, z) Q^1(x, dy) & \text{if } x \leq z \\ (1 - \beta)F(x, z) & \text{if } x > z \end{cases}$$

and

$$g(x, z) = \begin{cases} c(x, 1) + \beta \int_{(z, u]} G(y, z) Q^1(x, dy) & \text{if } x \leq z \\ (1 - \beta)G(x, z) & \text{if } x > z. \end{cases} \quad (80)$$

Since $c(x, 1) > 0$ for every x (see Assumption 2.1(i)), it follows immediately from (80) that $g(x, z) > 0$ for every x and z , and so PCL-indexability condition (PCLI1) holds.

Thus, the MP metric $m(x, z) \triangleq f(x, z)/g(x, z)$ is well defined, and so is the MP index, given by

$$m^*(x) = \frac{r(x, 1) + \beta \int_{(x, u]} F(y, x) Q^1(x, dy)}{c(x, 1) + \beta \int_{(x, u]} G(y, x) Q^1(x, dy)} = \frac{F(x, x^-)}{G(x, x^-)}, \quad x \in X,$$

where the second identity holds by Lemma 7.8(c).

Hence, if for the model of concern the MP index is shown to satisfy conditions (PCLI2, PCLI3), Theorem 2.1 ensures that the model is strongly threshold indexable with Gittins index m^* .

12.2 Optimal dynamic transmission over a noisy channel.

This is the model considered in Liu and Zhao [29, 30] and in Niño-Mora [38]. In the corresponding single-project subproblems, a user dynamically attempts to transmit packets over a noisy communication channel of Gilbert–Elliott type. In each discrete time period $t = 0, 1, \dots$, the channel is in one of two states: 1 (good), in which an attempted packet transmission will be successful, and 0, in which it will fail. Channel state transitions are Markovian, with probabilities p and q for the transitions $1 \rightarrow 0$ and $0 \rightarrow 1$, respectively. The channel state process autocorrelation $\rho \triangleq 1 - p - q$ is assumed to be positive. The user cannot observe the channel state, basing instead decisions on the *belief state* $X_t \in \mathbf{X} \triangleq [0, 1]$, giving the posterior probability that the channel state is 1, which is updated in a Bayesian fashion. Thus, if $X_t = x$ and the user takes the active action $A_t = 1$ (attempt to transmit) then the next belief state X_{t+1} will take the values $q + \rho$ and q with probabilities x and $1 - x$, respectively. If instead the user takes the passive action (do not attempt to transmit), then $X_{t+1} = q + \rho x$. A successful transmission yields a unit reward, and hence the reward function is $r(x, a) \triangleq ax$, while the resource consumption function is $c(x, a) \triangleq a$. The cost per transmission attempt is λ .

For a given discount factor β , the λ -price problem P_λ in (4) is to find an admissible transmission policy maximizing the expected total discounted net value of transmission attempts. To apply the present framework to analyze the model's indexability we must evaluate the performance metrics. For the z -policy, the reward and resource metrics are determined by the functional equations

$$\begin{aligned} F(x, z) &= \beta F(q + \rho x, z) 1_{\{x \leq z\}} + \{x + \beta x F(q + \rho, z) + \beta(1 - x)F(q, z)\} 1_{\{x > z\}}, \\ G(x, z) &= \beta G(q + \rho x, z) 1_{\{x \leq z\}} + \{1 + \beta x G(q + \rho, z) + \beta(1 - x)G(q, z)\} 1_{\{x > z\}}, \end{aligned}$$

and the corresponding marginal metrics are given by

$$\begin{aligned} f(x, z) &= x - \beta \{F(q + \rho x, z) - xF(q + \rho, z) - (1 - x)F(q, z)\}, \\ g(x, z) &= 1 - \beta \{G(q + \rho x, z) - xG(q + \rho, z) - (1 - x)G(q, z)\}. \end{aligned}$$

Provided $g(x, x) \neq 0$, the MP index is thus given by

$$m^*(x) = \frac{x - \beta \{F(q + \rho x, x) - xF(q + \rho, x) - (1 - x)F(q, x)\}}{1 - \beta \{G(q + \rho x, x) - xG(q + \rho, x) - (1 - x)G(q, x)\}}$$

The analysis below of performance metrics under threshold policies refers to the map $x \mapsto h(x) \triangleq q + \rho x$, and to the *forward iterates* $h_t(x)$, with $h_0(x) = x$ and $h_t(x) = h(h_{t-1}(x))$ for $t \geq 1$, which converge to $h_\infty \triangleq q/(1 - \rho)$ as $t \rightarrow \infty$. Note that $h_t(x) = h_\infty - (h_\infty - x)\rho^t$. We further refer to the *backward iterates* $h_{-t}(z) = h_\infty - (h_\infty - z)\rho^{-t}$ for $t \geq 1$. Four cases need be considered depending on the value of the threshold variable z , as outlined next.

Case I: $z < q$. In this case the belief state remains above threshold at times $t \geq 1$. The performance metrics have the evaluations

$$\begin{aligned} F(x, z) &= \frac{\beta(q + (1 - \beta)\rho x) + (1 - \beta)(1 - \beta\rho)x 1_{\{x > z\}}}{(1 - \beta)(1 - \beta\rho)}, & f(x, z) &= x \\ G(x, z) &= \frac{\beta + (1 - \beta)1_{\{x > z\}}}{1 - \beta}, & g(x, z) &= 1. \end{aligned}$$

Thus, condition (PCLI1) holds for every x when $z < q$, and the MP index is $m^*(x) = x$ for $x < q$.

Case II: $q \leq z < h_\infty$. Consider the subcase $h_{t-1}(q) \leq z < h_t(q)$ for $t \geq 1$. The performance metrics can be given in closed form in terms of $G_t(q) \triangleq G(q, h_{t-1}(q))$, $G_t(q + \rho) \triangleq G(q + \rho, h_{t-1}(q))$, $F_t(q) \triangleq F(q, h_{t-1}(q))$ and $F_t(q + \rho) \triangleq F(q + \rho, h_{t-1}(q))$, which are readily evaluated. Thus, for $x > z$,

$$\begin{aligned} F(x, z) &= x + \beta\{xF_t(q + \rho) + (1 - x)F_t(q)\} \\ f(x, z) &= x - \beta\{h(x) + \beta h(x)F_t(q + \rho) + \beta(1 - h(x))F_t(q) - xF_t(q + \rho) - (1 - x)F_t(q)\} \\ G(x, z) &= 1 + \beta\{xG_t(q + \rho) + (1 - x)G_t(q)\} \\ g(x, z) &= 1 - \beta\{1 + \beta h(x)G_t(q + \rho) + \beta(1 - h(x))G_t(q) - xG_t(q + \rho) - (1 - x)G_t(q)\}. \end{aligned}$$

For $x \leq z$, letting $s \leq t + 1$ be such that $h_{s-1}(x) \leq z < h_s(x)$, we have

$$\begin{aligned} F(x, z) &= \beta^s\{h_s(x) + \beta h_s(x)F_t(q + \rho) + \beta(1 - h_s(x))F_t(q)\} \\ f(x, z) &= x - \beta\{\beta^{s-1}[h_s(x) + \beta h_s(x)F_t(q + \rho) + \beta(1 - h_s(x))F_t(q)] - xF_t(q + \rho) - (1 - x)F_t(q)\} \\ G(x, z) &= \beta^s\{1 + \beta h_s(x)G_t(q + \rho) + \beta(1 - h_s(x))G_t(q)\} \\ g(x, z) &= 1 - \beta\{\beta^{s-1}[1 + \beta h_s(x)G_t(q + \rho) + \beta(1 - h_s(x))G_t(q)] - xG_t(q + \rho) - (1 - x)G_t(q)\}. \end{aligned}$$

Hence, for $q \leq x < h_\infty$, and provided $g(x, x) > 0$, the MP index is given by

$$m^*(x) = \frac{x - \beta\{h(x) + \beta h(x)F_t(q + \rho) + \beta(1 - h(x))F_t(q) - xF_t(q + \rho) - (1 - x)F_t(q)\}}{1 - \beta\{1 + \beta h(x)G_t(q + \rho) + \beta(1 - h(x))G_t(q) - xG_t(q + \rho) - (1 - x)G_t(q)\}}.$$

Case III: $h_\infty \leq z < q + \rho$. In this case, the performance metrics have the evaluations

$$\begin{aligned} F(x, z) &= \frac{x}{1 - \beta(q + \rho)}1_{\{x > z\}} \\ f(x, z) &= \frac{x}{1 - \beta(q + \rho)}1_{\{h(x) \leq z\}} + \frac{(1 - \beta\rho)x - \beta q}{1 - \beta(q + \rho)}1_{\{h(x) > z\}} \\ G(x, z) &= \frac{1 - \beta(q + \rho - x)}{1 - \beta(q + \rho)}1_{\{x > z\}} \\ g(x, z) &= \frac{1 - \beta(q + \rho - x)}{1 - \beta(q + \rho)}1_{\{h(x) \leq z\}} + \left(1 - \beta + \beta \frac{(1 - \beta\rho)x - \beta q}{1 - \beta(q + \rho)}\right)1_{\{h(x) > z\}}. \end{aligned}$$

Hence, for $h_\infty \leq x < q + \rho$, as $h(x) \leq x$, the MP index is given by

$$m^*(x) = \frac{x}{1 - \beta(q + \rho - x)}.$$

Case IV: $z \geq q + \rho$. Finally, in this case we have

$$\begin{aligned} F(x, z) &= x1_{\{x > z\}}, & f(x, z) &= x \\ G(x, z) &= 1_{\{x > z\}}, & g(x, z) &= 1, \end{aligned}$$

and therefore, for $x \geq q + \rho$, the MP index is given by $m^*(x) = x$.

Proposition 12.1. *The above optimal channel transmission model is PCL-indexable.*

Proof. Proof. We provide a proof outline due to space limitations and to avoid tedious detail. Consider (PCLI1). One can show the stronger condition $g(x, z) \geq 1 - \beta$ for every x and z , which is trivial in cases I, III and IV. In case III, for $h_{t-1}(q) < z < h_t(q)$, $g(x, z)$ is piecewise linear and càglàd in x with $t + 1$ pieces, with each piece being increasing, which follows from $G_t(q + \rho) > G_t(q)$ and

$$\frac{\partial}{\partial x} g(x, z) = \begin{cases} \beta(1 - (\beta\rho)^{t+1})(G_t(q + \rho) - G_t(q)), & \text{for } 0 < x < h_{-t}(z) \\ \beta(1 - (\beta\rho)^s)(G_t(q + \rho) - G_t(q)), & \text{for } h_{-s}(z) < x < h_{-s+1}(z), s = 2, \dots, t \\ \beta(1 - \beta\rho)(G_t(q + \rho) - G_t(q)), & \text{for } h_{-1}(z) < x < 1. \end{cases}$$

Hence, $g(x, z) > g(0^+, z)$ for $0 < x \leq h_{-t}(z)$, $g(x, z) > g(h_{-s}(z)^+, z)$ for $h_{-s}(z) < x \leq h_{-s+1}(z)$ with $2 \leq s \leq t$, and $g(x, z) > g(h_{-1}(z)^+, z)$ for $h_{-1}(z) < x \leq 1$. Now, it can be verified that $g(0^+, z) = 1$, and, using that $(\partial/\partial z)g(h_{-s}(z)^+, z) = \beta(\rho^{-s} - \beta^s)(G_t(q + \rho) - G_t(q)) > 0$, $g(h_{-s}(z)^+, z) > g(h_{t-s-1}(q)^+, h_{t-1}(q))$ for $s = 1, \dots, t$. One can then show through algebraic and calculus arguments that $g(h_{t-s-1}(q)^+, h_{t-1}(q)) \geq 1 - \beta$ for each such s , and thus establish satisfaction of (PCLI1).

As for condition (PCLI2), it is also easily verified in cases I, III and IV, viz., for $x \in [0, q) \cup [h_\infty, 1]$. Its verification in case II involves tedious algebraic and calculus arguments.

Regarding (PCLI3), we use the sufficient conditions given in §B.2. Thus, taking

$$\Omega(x) \triangleq \{h_t(x) : t \geq 0\} \cup \{h_t(q) : t \geq 0\} \cup \{h_\infty, q, q + \rho\},$$

Assumption B.3 holds, whence Lemma B.5 and Proposition B.2 yield (PCLI3). \square

Thus, Theorem 2.1 yields that the model is strongly threshold indexable with Whittle index m^* . While the present framework applies to the discounted criterion, note that the index m^* above converges as $\beta \rightarrow 1$ to a limiting index, which is precisely the long-run average Whittle index derived in Liu and Zhao [30].

13 Concluding remarks.

This paper has presented sufficient conditions for indexability of general restless bandits in a real state discrete time discounted setting, rigorously demonstrating their validity. This extends the approach previously developed by the author for discrete state projects in work referred to above. Among the issues raised by the present work we highlight that further work is required to establish the applicability of the proposed conditions to a variety of relevant model classes, and to identify properties satisfied by model primitives that imply satisfaction of the PCL-indexability conditions.

Appendices

A Geometric and economic interpretations of the MP index.

This appendix gives geometric and economic interpretations of the MP index, showing in particular that it can be characterized as a *resource shadow price* under PCL-indexability, extending corresponding results in Niño-Mora [35, 36] for discrete state projects. Consider the *achievable resource-reward performance region* $R_{GF}(p)$ with initial state $X_0 \sim p \in \mathbb{P}_w(\mathbf{X})$, which is the region in the resource-reward γ - ϕ plane spanned by points $(G(p, \pi), F(p, \pi))$, i.e.,

$$R_{GF}(p) \triangleq \{(G(p, \pi), F(p, \pi)) : \pi \in \Pi\} = \{(\gamma, \phi) : (G(p, \pi), F(p, \pi)) = (\gamma, \phi) \text{ for some } \pi \in \Pi\}. \quad (81)$$

When starting from state x , we write $R_{GF}(x)$.

Lemma A.1. *The region $R_{GF}(p)$ is compact and convex.*

Proof. Proof. This follows from the results reviewed in §4 and the representation of $R_{GF}(p)$ as

$$R_{GF}(p) = \{(\langle \mu, c \rangle, \langle \mu, r \rangle) : \mu \in \mathcal{M}_p\},$$

since \mathcal{M}_p is compact and convex, and both $\langle \cdot, c \rangle$ and $\langle \cdot, r \rangle$ are continuous linear functionals. Note that the result that $R_{GF}(p)$ is bounded also follows from (8). \square

We will be interested in characterizing the *upper resource-reward boundary* $\text{ub}(R_{GF}(p))$, which consists of all achievable resource-reward performance points attaining the maximum reward performance for their resource usage performance, i.e.,

$$\text{ub}(R_{GF}(p)) = \{(\gamma, \phi) \in R_{GF}(p) : \phi = \max\{\tilde{\phi} : (\gamma, \tilde{\phi}) \in R_{GF}(p)\}\}.$$

For such a purpose, we introduce a parametric family of constrained MDPs as follows. First, note that it is immediate from Lemma A.1 that the *achievable resource (usage) performance region* $R_G(p) \triangleq \{G(p, \pi) : \pi \in \Pi\}$ is a compact interval. Now, for each achievable resource performance $\gamma \in R_G(p)$ consider the γ -*resource problem*

$$\begin{aligned} C_p(\gamma) : \quad & \text{maximize } F(p, \pi) \\ & \text{subject to: } G(p, \pi) = \gamma, \pi \in \Pi. \end{aligned} \tag{82}$$

Let $\Phi_p(\gamma)$ denote the optimal value of problem $C_p(\gamma)$. Note that, below, \mathring{S} denotes the interior of a set S .

Remark A.1. (i) In light of Lemma A.1, the upper boundary $\text{ub}(R_{GF}(p))$ is the graph of the function $\Phi_p : R_G(p) \rightarrow \mathbb{R}$, which is concave (and hence continuous).

(ii) By standard results in nonsmooth convex analysis (cf. Rockafellar [45, Ch. 23]) Φ_p has finite left and right derivatives $\Phi_p'^-(\gamma)$ and $\Phi_p'^+(\gamma)$ at each $\gamma \in \mathring{R}_G(p)$, with $\Phi_p'^-(\gamma) \geq \Phi_p'^+(\gamma)$, which are, respectively, the *left* and the *right shadow prices* of the resource constraint $G(p, \pi) = \gamma$ in $C_p(\gamma)$, or, in economic terms, the *left* and the *right MP of the resource*. If $\Phi_p'^-(\gamma) = \Phi_p'^+(\gamma)$ then Φ_p is differentiable at γ , and we write its derivative as $\Phi_p'(\gamma)$.

(iii) Note that a price $\lambda \in \mathbb{R}$ is a *supergradient* of Φ_p at $\gamma_0 \in R_G(p)$ if

$$\Phi_p(\gamma) \leq \Phi_p(\gamma_0) + \lambda(\gamma - \gamma_0), \quad \gamma \in R_G(p).$$

i.e., if $\{(\gamma, \phi) \in \mathbb{R}^2 : \phi - \lambda\gamma = \Phi_p(\gamma_0) - \lambda\gamma_0\}$ is a nonvertical supporting line to the convex hypograph $\text{hyp}(\Phi_p)$ of Φ_p at $(\gamma_0, \Phi_p(\gamma_0))$. The set of supergradients of Φ_p at $\gamma_0 \in \mathring{R}_G(p)$ is $[\Phi_p'^+(\gamma_0), \Phi_p'^-(\gamma_0)]$.

The following result, which assumes PCL-indexability condition (PCLI1), characterizes in its part (a) the resource performance region $R_G(p)$, whereas its part (b) establishes that every point of $R_G(p)$ can be achieved through a randomized threshold policy.

Lemma A.2. *Let (PCLI1) hold. Then*

- (a) $R_G(p) = [G(p, \infty), G(p, -\infty)]$;
- (b) *for every $\gamma \in R_G(p)$ there exist $z \in \overline{\mathbb{R}}$ and $\alpha \in [0, 1]$ such that $G(p, z) \leq \gamma \leq G(p, z^-)$ and $G(p, \alpha z + (1 - \alpha)z^-) = \gamma$.*

Proof. Proof. (a) This part follows by combining Lemma A.1, which yields that $R_G(p)$ is a compact interval, with Lemma 6.1, which ensures satisfaction of the PCLs in Definition 6.1 under (PCLI1). Thus, for every admissible π , PCL(a) along with (PCLI1) implies $G(p, \pi) \leq G(p, -\infty)$, whereas PCL(b.1) gives $G(p, \pi) \geq G(p, \infty)$.

(b) Consider the strict upper level set $U(\gamma) \triangleq \{z \in \mathbb{R} : G(p, z) > \gamma\}$ and let $\bar{z}(\gamma) \triangleq \sup U(\gamma)$, where the “sup” over the empty set is taken to be $-\infty$. If $U(\gamma) = \mathbb{R}$ then $\bar{z}(\gamma) = \infty$, and $\gamma \in R_G(p)$, part (a) and Lemma 7.2(c) yield $\gamma = G(p, \infty)$. If $U(\gamma) = \emptyset$ then $\bar{z}(\gamma) = -\infty$, and $\gamma \in R_G(p)$, part (a) and Lemma 7.2(b) yield $\gamma = G(p, -\infty)$. Otherwise, $\bar{z}(\gamma)$ is finite and, since $G(p, \cdot)$ is càdlàg by Lemma 7.2 and nonincreasing by Lemma 7.3(a), it follows that $G(p, \bar{z}(\gamma)) \leq \gamma \leq G(p, \bar{z}(\gamma)^-)$. So, in any case, let $z = \bar{z}(\gamma)$ and pick $\alpha \in [0, 1]$ such that $\alpha G(p, z) + (1 - \alpha)G(p, z^-) = \gamma$. Then (see Remark 2.3(ii)) the randomized threshold policy $\alpha z + (1 - \alpha)z^-$ satisfies $(G(p, \alpha z + (1 - \alpha)z^-), F(p, \alpha z + (1 - \alpha)z^-)) = (\gamma, \alpha F(p, z) + (1 - \alpha)F(p, z^-))$. \square

Consider now the subregion $R_{GF}^{\mathcal{Z}}(p)$ of $R_{GF}(p)$ spanned by threshold policies,

$$R_{GF}^{\mathcal{Z}}(p) \triangleq \{(G(p, \pi), F(p, \pi)) : \pi = z \text{ or } \pi = z^- \text{ for some } z \in \overline{\mathbb{R}}\}. \quad (83)$$

Denote by $\text{co}(R_{GF}^{\mathcal{Z}}(p))$ its convex hull and by $\text{uh}(R_{GF}^{\mathcal{Z}}(p)) \triangleq \text{ub}(\text{co}(R_{GF}^{\mathcal{Z}}(p)))$ its *upper hull* (the upper boundary of its convex hull), i.e.,

$$\text{uh}(R_{GF}^{\mathcal{Z}}(p)) = \{(\gamma, \phi) \in \text{co}(R_{GF}^{\mathcal{Z}}(p)) : \phi = \sup\{\tilde{\phi} : (\gamma, \tilde{\phi}) \in \text{co}(R_{GF}^{\mathcal{Z}}(p))\}\}.$$

The following result shows that, under PCL-indexability, the upper boundary $\text{ub}(R_{GF}(p))$ is completely characterized by the performance points of threshold policies.

Lemma A.3. *Let the project be PCL-indexable. Then*

- (a) *for every $\gamma \in R_G(p)$, $z \in \overline{\mathbb{R}}$ and $\alpha \in [0, 1]$ with $G(p, \alpha z + (1 - \alpha)z^-) = \gamma$, the randomized threshold policy $\alpha z + (1 - \alpha)z^-$ is $C_p(\gamma)$ -optimal, and so $\Phi_p(\gamma) = F(p, \alpha z + (1 - \alpha)z^-)$;*
- (b) $R_{GF}^{\mathcal{Z}}(p) \subseteq \text{ub}(R_{GF}(p))$;
- (c) $\text{ub}(R_{GF}(p)) = \text{uh}(R_{GF}^{\mathcal{Z}}(p))$.

Proof. Proof. (a) Let z and α be as in Lemma A.2(b), so $G(p, \alpha z + (1 - \alpha)z^-) = \gamma$. On the one hand, we have $\Phi_p(\gamma) \triangleq \max\{F(p, \pi) : G(p, \pi) = \gamma, \pi \in \Pi\} \geq F(p, \alpha z + (1 - \alpha)z^-)$. On the other hand, if $m^*(z) < \infty$, Theorem 2.1 ensures that both the z -policy and the z^- -policy are P_λ -optimal (see (4)) for $\lambda = m^*(z)$. Thus, $F(p, \pi) - m^*(z)G(p, \pi) \leq F(p, z) - m^*(z)G(p, z) = F(p, z^-) - m^*(z)G(p, z^-)$ for every $\pi \in \Pi$, whence (see Remark 2.3(ii)) $F(p, \pi) - m^*(z)G(p, \pi) \leq F(p, \alpha z + (1 - \alpha)z^-) - m^*(z)\gamma$ for every $\pi \in \Pi$, which yields $\Phi_p(\gamma) \leq F(p, \alpha z + (1 - \alpha)z^-)$. Therefore, $\Phi_p(\gamma) = F(p, \alpha z + (1 - \alpha)z^-)$.

As for the case $m^*(z) = \infty$, it must then be $z = u = \infty$, so $\alpha z + (1 - \alpha)z^-$ is the ∞ -policy (“never active”). Now, Lemma 5.2(b) and $G(p, \infty) = \gamma$ yield $G(p, \pi) = \gamma + \int g(y, \infty) \tilde{\mu}_p^{\pi, 1}(dy)$ for every admissible π . Using (PCLI1), it follows that $\tilde{\mu}_p^{\pi, 1}(\mathbf{X}) = 0$ for every $C_p(\gamma)$ -feasible π . Combining this result

with Lemma 5.2(a) yields that, for every $C_p(\gamma)$ -feasible π , $F(p, \pi) = F(p, \infty) + \int f(y, \infty) \tilde{\mu}_p^{\pi, 1}(dy) = F(p, \infty)$, whence $\Phi_p(\gamma) = F(p, \infty)$.

(b) This part is an immediate consequence of part (a).

(c) Since $\text{ub}(\mathbf{R}_{GF}(p))$ is the graph of a concave function, (b) implies $\text{uh}(\mathbf{R}_{GF}^{\mathcal{Z}}(p)) \subseteq \text{ub}(\mathbf{R}_{GF}(p))$.

Let $(\gamma, \phi) \in \text{ub}(\mathbf{R}_{GF}(p))$. Then by part (a) there exist z and α such that $(\gamma, \phi) = \alpha(G(p, z), F(p, z)) + (1 - \alpha)(G(p, z^-), F(p, z^-))$, whence $(\gamma, \phi) \in \text{co}(\mathbf{R}_{GF}^{\mathcal{Z}}(p))$. Using that $\text{co}(\mathbf{R}_{GF}^{\mathcal{Z}}(p)) \subseteq \mathbf{R}_{GF}(p)$, this immediately yields $(\gamma, \phi) \in \text{uh}(\mathbf{R}_{GF}^{\mathcal{Z}}(p))$.

On the other hand, let $(\gamma, \phi) \in \text{uh}(\mathbf{R}_{GF}^{\mathcal{Z}}(p))$. Lemma A.1 then yields $(\gamma, \phi) \in \mathbf{R}_{GF}(p)$. Now, by Lemma A.2(b) there exist z and α such that $G(p, \alpha z + (1 - \alpha)z^-) = \gamma$ and hence, by part (a), $\Phi_p(\gamma) = F(p, \alpha z + (1 - \alpha)z^-)$. This immediately yields $\phi = \Phi_p(\gamma)$, whence $(\gamma, \phi) \in \text{ub}(\mathbf{R}_{GF}(p))$. \square

Proposition A.1 (MP index as shadow price). *Let the project be PCL-indexable. If p has full support \mathbf{X} then Φ_p is differentiable on $\mathring{\mathbf{R}}_G(p)$, with $\Phi'_p(\gamma) = m^*(z)$ if $G(p, z) \leq \gamma \leq G(p, z^-)$.*

Proof. Proof. Let $\gamma \in \mathring{\mathbf{R}}_G(p)$, i.e., by Lemma A.2(a), $\gamma \in (G(p, \infty), G(p, -\infty))$. By Lemma A.2(b), there exist $z \in \mathbb{R}$ and $\alpha \in [0, 1]$ such that $G(p, \alpha z + (1 - \alpha)z^-) = \gamma$. Since $G(p, \cdot)$ is nonincreasing by Lemma 7.3(a), and (see Remark 2.3(ii)) $G(p, \alpha z + (1 - \alpha)z^-) = \alpha G(p, z) + (1 - \alpha)G(p, z^-)$, we have $G(p, z) \leq \gamma \leq G(p, z^-)$. Consider first the case $G(p, z) < \gamma < G(p, z^-)$, which implies $z \in \mathbf{X}$. Then the unique α satisfying the above is $\alpha(\gamma) \triangleq (G(p, z^-) - \gamma)/(G(p, z^-) - G(p, z))$. Thus, by Lemma A.3(a), $\Phi_p(\gamma) = F(p, \alpha(\gamma)z + (1 - \alpha(\gamma))z^-) = \alpha(\gamma)F(p, z) + (1 - \alpha(\gamma))F(p, z^-)$, and hence Φ_p is linear on $(G(p, z), G(p, z^-))$, with $\Phi'_p(\gamma) = (F(p, z^-) - F(p, z))/(G(p, z^-) - G(p, z)) = m^*(z)$ for $\gamma \in (G(p, z), G(p, z^-))$, where we have used Lemma 7.8(a).

Consider now the case $\gamma = G(p, z) < G(p, z^-)$. Then $z \in \mathbf{X} \setminus \{u\}$ and, arguing along the lines of the previous case yields that the right derivative of Φ_p at γ is given by $\Phi_p'^+(\gamma) = m^*(z)$. To evaluate the left derivative $\Phi_p'^-$ (which exists and is finite; see Remark A.1(iii)) consider a decreasing threshold sequence $\{z_n\}$ in \mathbf{X} converging to z . Letting $\gamma_n \triangleq G(p, z_n)$, by Lemmas 7.3(b) and 7.2 the sequence $\{\gamma_n\}$ is increasing and converges to $G(p, z) = \gamma$. Now, Lemma A.3(a) and Proposition 7.2 yield

$$\Phi_p'^-(\gamma) = \lim_{n \rightarrow \infty} \frac{\Phi_p(\gamma) - \Phi_p(\gamma_n)}{\gamma - \gamma_n} = \lim_{n \rightarrow \infty} \frac{F(p, z) - F(p, z_n)}{G(p, z) - G(p, z_n)} = m^*(z).$$

In the case $G(p, z) < G(p, z^-) = \gamma$, $z \in \mathbf{X} \setminus \{\ell\}$ and, arguing along the lines of the previous case yields that the left derivative of Φ_p at γ is given by $\Phi_p'^-(\gamma) = m^*(z)$. To evaluate the right derivative $\Phi_p'^+$ consider an increasing threshold sequence $\{z_n\}$ in \mathbf{X} converging to z . Letting $\gamma_n \triangleq G(p, z_n)$ we further obtain that the sequence $\{\gamma_n\}$ is decreasing and converges to $G(p, z^-) = \gamma$ and, furthermore,

$$\Phi_p'^+(\gamma) = \lim_{n \rightarrow \infty} \frac{\Phi_p(\gamma_n) - \Phi_p(\gamma)}{\gamma_n - \gamma} = \lim_{n \rightarrow \infty} \frac{F(p, z_n) - F(p, z^-)}{G(p, z_n) - G(p, z^-)} = m^*(z).$$

Finally, in the case $\gamma = G(p, z) = G(p, z^-)$ the result follows along the same lines. \square

Figure 1 shows the upper resource-reward boundary $\text{ub}(\mathbf{R}_{GF}(p))$ for the model in §12.2 when $X_0 \sim \text{Uniform}[0, 1]$. Note that in the plot $F(p, z)$ and $G(p, z)$ are written as $F(z)$ and $G(z)$.

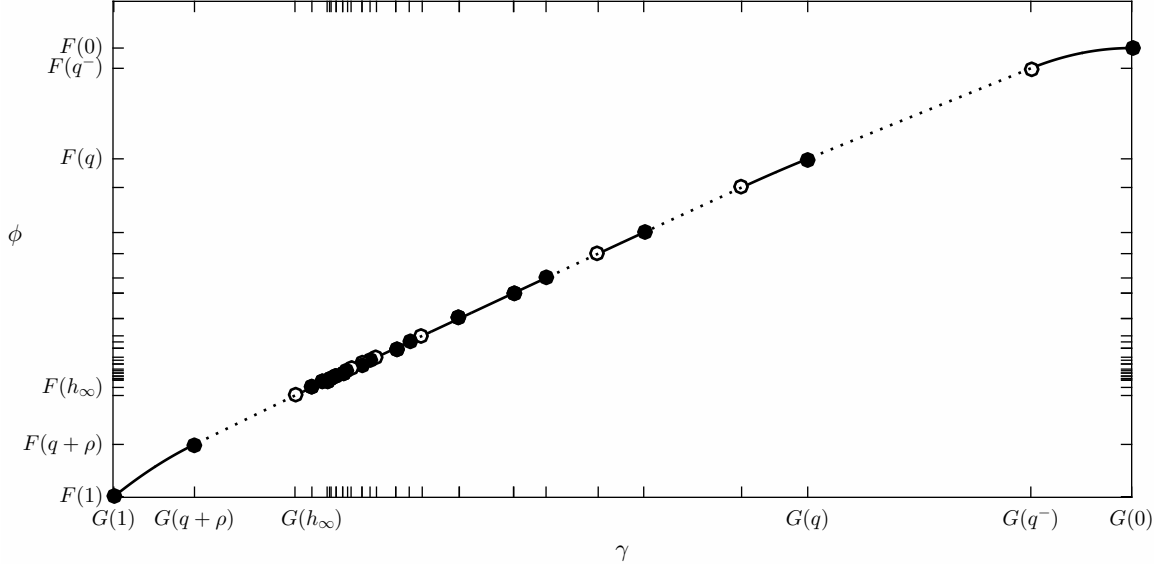


Figure 1: The upper resource-reward boundary $\text{ub}(\mathbf{R}_{GF}(p))$ for the model in §12.2.

B Checking condition (PCLI3).

Since a direct verification of condition (PCLI3) in Definition 2.4 for a particular model may be hard or tedious, this appendix gives conditions under which (PCLI3) can be established by simpler means. The main results are given in Propositions B.1 and B.2.

B.1 Piecewise differentiable $F(x, \cdot)$ and $G(x, \cdot)$.

We start with the case where the reward and the resource metrics $F(x, z)$ and $G(x, z)$ are piecewise differentiable with respect to the threshold variable z . We denote by $F'_2(x, z)$ and $G'_2(x, z)$ the partial derivatives of $F(x, z)$ and $G(x, z)$ with respect to z . Recall that \mathring{S} denotes the interior of a set S .

Assumption B.1. *For each state x there exists a countable partition $\{S_i(x) : i \in N(x)\}$ of \mathbb{R} with each $S_i(x)$ being either a left semiclosed interval or a singleton set, such that:*

- (i) $F(x, \cdot) = \sum_{i \in N(x)} F_i(x, \cdot) 1_{S_i(x)}$ and $G(x, \cdot) = \sum_{i \in N(x)} G_i(x, \cdot) 1_{S_i(x)}$, with $F_i(x, \cdot)$ and $G_i(x, \cdot)$ being both differentiable on $\mathring{S}_i(x)$.
- (ii) $F'_2(x, z) = m^*(z) G'_2(x, z)$ for every $i \in N(x)$ and $z \in \mathring{S}_i(x)$.

We need two preliminary results. Let $\text{TV}_{F(x, \cdot)}(I)$ and $\text{TV}_{G(x, \cdot)}(I)$ denote the *total variation* of $F(x, \cdot)$ and $G(x, \cdot)$ on an interval I , respectively, and let $D(x) \triangleq \mathbb{R} \setminus \cup_{i \in N(x)} \mathring{S}_i(x)$ be the countable set of interval endpoints for the partition $\{S_i(x) : i \in N(x)\}$. Recall (see §7.4) that $\mathbb{V}(I)$ denotes the class of functions of bounded variation on the set I .

Lemma B.1. *Let (PCLI1, PCLI2) and Assumption B.1 hold. Let x be a state and I a finite open interval where both $F(x, \cdot)$ and $G(x, \cdot)$ are differentiable. Then $F(x, \cdot) \in \mathbb{V}(I)$ and*

$$\int_I F(x, dz) = \int_I m^*(z) G(x, dz).$$

Proof. Proof. Since $G(x, \cdot)$ is bounded nonincreasing by Lemma 7.3(a) and m^* is continuous, by standard results $\int_I |m^*|(z) G(x, dz)$ exists and is finite. Furthermore, using Assumption B.1(ii),

$$\int_I |m^*|(z) G(x, dz) = \int_I |m^*|(z) G'_2(x, z) dz = - \int_I |F'_2|(x, z) dz = -\text{TV}_{F(x, \cdot)}(I)$$

(cf. Carter and van Brunt [8, The. 6.1.7]). Hence, $F(x, \cdot) \in \mathbb{V}(I)$ and

$$\int_I m^*(z) G(x, dz) = \int_I m^*(z) G'_2(x, z) dz = \int_I F'_2(x, z) dz = \int_I F(x, dz).$$

□

In the next result, $\mathbb{V}_{\text{loc}}(\mathbb{R})$ denotes the class of functions of *locally bounded variation* on \mathbb{R} , i.e., having bounded variation on every finite interval.

Lemma B.2. *Under (PCLI1, PCLI2) and Assumption B.1, $F(x, \cdot) \in \mathbb{V}_{\text{loc}}(\mathbb{R})$.*

Proof. Proof. Let $I = (z_1, z_2)$ with $-\infty < z_1 < z_2 < \infty$ and $K(I) \triangleq \max\{|m^*|(z) : z \in [z_1, z_2]\}$, which is finite by continuity of m^* . The total variation $\text{TV}_{F(x, \cdot)}(I)$ of $F(x, \cdot)$ on I is finite, since

$$\begin{aligned} \text{TV}_{F(x, \cdot)}(I) &= \sum_{z \in D(x) \cap I} |\Delta_2 F(x, z)| + \sum_{i: \dot{S}_i(x) \cap I \neq \emptyset} \int_{\dot{S}_i(x) \cap I} |F'_2|(x, z) dz \\ &= \sum_{z \in D(x) \cap I} |m^*|(z) |\Delta_2 G(x, z)| + \sum_{i: \dot{S}_i(x) \cap I \neq \emptyset} \int_{\dot{S}_i(x) \cap I} |m^*|(z) |G'_2|(x, z) dz \\ &\leq K(I) \left\{ \sum_{z \in D(x) \cap I} |\Delta_2 G(x, z)| + \int_{i: \dot{S}_i(x) \cap I \neq \emptyset} |G'_2|(x, z) dz \right\} \\ &= K(I) \text{TV}_{G(x, \cdot)}(I) \leq K(I) \text{TV}_{G(x, \cdot)}(\mathbb{R}) < \infty, \end{aligned}$$

where we have used Assumption B.1 and Lemmas 7.8(a), B.1 and 7.9. □

We can now give the main result of this section.

Proposition B.1. *Under (PCLI1, PCLI2) and Assumption B.1, (PCLI3) holds.*

Proof. Proof. Let $-\infty < z_1 < z_2 < \infty$. We have

$$\begin{aligned} F(x, z_2) - F(x, z_1) &= \int_{(z_1, z_2]} F(x, dz) = \Delta_2 F(x, z_2) + \int_{(z_1, z_2)} F(x, dz) \\ &= \sum_{z \in D(x) \cap (z_1, z_2]} \Delta_2 F(x, z) + \sum_{i: \dot{S}_i(x) \cap (z_1, z_2) \neq \emptyset} \int_{\dot{S}_i(x) \cap (z_1, z_2)} F'_2(x, z) dz \\ &= \sum_{z \in D(x) \cap (z_1, z_2]} m^*(z) \Delta_2 G(x, z) + \sum_{i: \dot{S}_i(x) \cap (z_1, z_2) \neq \emptyset} \int_{\dot{S}_i(x) \cap (z_1, z_2)} m^*(z) G'_2(x, z) dz \\ &= \int_{(z_1, z_2]} m^*(z) G(x, dz), \end{aligned}$$

where we have used Lemma B.2, Assumption B.1 and Lemmas 7.8(a), B.1 and 7.4. □

B.2 Piecewise constant $G(x, \cdot)$.

We next consider the case where the resource metric $G(x, z)$ is piecewise constant in the threshold variable z .

Assumption B.2. For each state x , $G(x, \cdot) = \sum_{i \in N(x)} G_i(x) 1_{S_i(x)}$, where $\{S_i(x) : i \in N(x)\}$ is a countable partition of \mathbb{R} with each $S_i(x)$ being either a left semiclosed interval or a singleton set.

Note that, unlike Assumption B.1, Assumption B.2 does not refer to the reward metric $F(x, z)$, so it is not evident that this case reduces to that considered in §B.1. Yet, we will show that Assumption B.2 does imply Assumption B.1.

The following auxiliary result can be used to verify satisfaction of Assumption B.2 in a given model. The result gives, in parts (a, c), conditions characterizing when $G(x, \cdot)$ is constant on an interval, in terms of the project state dynamics. As for parts (b, d), they ensure that under Assumption B.2 the function $F(x, \cdot)$ is constant over the same intervals that $G(x, \cdot)$, being hence of the form $F(x, \cdot) = \sum_{i \in N(x)} F_i(x) 1_{S_i(x)}$. For $S \in \mathcal{B}(\mathbb{R})$, we denote by $T_S \triangleq \min\{t \geq 0 : X_t \in S\}$ the first hitting time of the set S by the project state process, with $T_S \triangleq \infty$ if S is never hit.

Lemma B.3. Let (PCLI1) hold. For every state x and finite $z_1 < z_2$,

- (a) $G(x, \cdot)$ is constant on $[z_1, z_2]$ iff $\mathbf{P}_x^{z_1}\{T_{(z_1, z_2]} < \infty\} = 0$;
- (b) if $G(x, \cdot)$ is constant on $[z_1, z_2]$, then so is $F(x, \cdot)$;
- (c) $G(x, \cdot)$ is constant on $[z_1, z_2)$ iff $\mathbf{P}_x^{z_1}\{T_{(z_1, z_2)} < \infty\} = 0$;
- (d) if $G(x, \cdot)$ is constant on $[z_1, z_2)$, then so is $F(x, \cdot)$.

Proof. Proof. (a) Since $G(x, \cdot)$ is nonincreasing on \mathbb{R} by Lemma 7.3(a), it is constant on $[z_1, z_2]$ iff $G(x, z_1) = G(x, z_2)$, which occurs, by Lemma 7.1(b), iff

$$\int_{(z_1, z_2]} g(y, z_2) \tilde{\mu}_x^{z_1}(dy) = 0. \quad (84)$$

Now, under (PCLI1), (84) holds iff $\tilde{\mu}_x^{z_1}(z_1, z_2] = 0$, which occurs iff $\mathbf{P}_x^{z_1}\{T_{(z_1, z_2]} < \infty\} = 0$.

(b) If $G(x, \cdot)$ is constant on $[z_1, z_2]$ then $\tilde{\mu}_x^{z_1}(z_1, z_2] = 0$ which, along with Lemma 7.1(a), yields

$$F(x, z_1) - F(x, z) = \int_{(z_1, z]} f(y, z) \tilde{\mu}_x^{z_1}(dy) = 0, \quad z \in (z_1, z_2].$$

Hence, $F(x, \cdot)$ is also constant on $[z_1, z_2]$.

(c) Let $\{z^n\}_{n=0}^\infty$ be increasing with $z_1 < z^n < z_2$ and $\lim_{n \rightarrow \infty} z^n = z_2$. Then $G(x, \cdot)$ is constant on $[z_1, z_2)$ iff $G(x, \cdot)$ is constant on $[z_1, z^n]$ for each n , which by part (a) occurs iff $\tilde{\mu}_x^{z_1}(z_1, z^n] = 0$ for each n . By continuity from below, such will be the case iff $\tilde{\mu}_x^{z_1}(z_1, z_2) = 0$, i.e., iff $\mathbf{P}_x^{z_1}\{T_{(z_1, z_2)} < \infty\} = 0$.

(d) The proof of this part follows along the lines of that for part (b). \square

Note that, e.g., part (c) says that $G(x, \cdot)$ is constant on $[z_1, z_2)$ iff the project state never hits the interval (z_1, z_2) under the z_1 -policy starting from x .

Lemma B.4. Assumption B.2 implies Assumption B.1.

Proof. Proof. The result follows immediately from Lemma B.3(d). \square

Assumption B.3. For each state x there is a countable closed subset of states $\Omega(x)$, giving the endpoints of a countable disjoint partition of \mathbb{R} into open intervals, such that $\mathbb{P}_x^z\{X_t \in D(x)\} = 1$ for every time t and threshold z .

Lemma B.5. Assumption B.3 implies Assumption B.2.

Proof. Proof. The result follows immediately from Lemma B.3(c). \square

We thus obtain the following result.

Proposition B.2. Under (PCLI1, PCLI2) and Assumption B.2, (PCLI3) holds.

Proof. Proof. The result follows directly from Lemmas B.1 and B.4. \square

C Performance metrics and index computation.

Motivated by the fact that it will often be the case that, for the model at hand, the project metrics and index cannot be evaluated in closed form, this appendix gives recursions for approximately computing them.

For a state x , threshold z and horizon $k = 0, 1, \dots$, consider the k -horizon performance metrics

$$F_k(x, z) \triangleq \mathbb{E}_x^z \left[\sum_{t=0}^k \beta^t r(X_t, A_t) \right] \quad \text{and} \quad G_k(x, z) \triangleq \mathbb{E}_x^z \left[\sum_{t=0}^k \beta^t c(X_t, A_t) \right]. \quad (85)$$

Clearly, the function sequences $\{F_k(\cdot, z)\}_{k=0}^\infty$ and $\{G_k(\cdot, z)\}_{k=0}^\infty$ are determined by the following *value iteration* recursions: $F_0(x, z) \triangleq r(x, 1_{\{x > z\}})$, $G_0(x, z) \triangleq c(x, 1_{\{x > z\}})$ and, for $k = 0, 1, \dots$,

$$F_{k+1}(x, z) \triangleq r(x, 1_{\{x > z\}}) + \beta \int F_k(y, z) Q^{1_{\{x > z\}}}(x, dy), \quad (86)$$

$$G_{k+1}(x, z) \triangleq c(x, 1_{\{x > z\}}) + \beta \int G_k(y, z) Q^{1_{\{x > z\}}}(x, dy). \quad (87)$$

Consider further the k -horizon marginal metrics $f_k(x, z) \triangleq \Delta_{a=0}^{a=1} F_k(x, \langle a, z \rangle)$ and $g_k(x, z) \triangleq \Delta_{a=0}^{a=1} G_k(x, \langle a, z \rangle)$, which can be computed by $f_0(x, z) = \Delta_{a=0}^{a=1} r(x, a)$, $g_0(x, z) = \Delta_{a=0}^{a=1} c(x, a)$, and

$$f_{k+1}(x, z) = \Delta_{a=0}^{a=1} \{r(x, a) + \beta \int F_k(y, z) Q^a(x, dy)\} \quad (88)$$

$$g_{k+1}(x, z) = \Delta_{a=0}^{a=1} \{c(x, a) + \beta \int G_k(y, z) Q^a(x, dy)\} \quad (89)$$

Finally, consider the k -horizon MP metric $m_k(x, z) \triangleq f_k(x, z)/g_k(x, z)$ and the k -horizon MP index $\psi_k^{\text{MP}}(x) \triangleq m_k(x, x)$, which are defined when $g_k(x, z) \neq 0$ and $g_k(x, x) \neq 0$, respectively.

The following result shows, in parts (a, c), that the functions $F_k(\cdot, z)$, $G_k(\cdot, z)$, $f_k(\cdot, z)$ and $g_k(\cdot, z)$ converge in the w -norm $\|\cdot\|_w$ (see (5) and Remark 2.1(ii)) at least linearly with rate γ to their infinite-horizon counterparts. Parts (b, d) further bound the w -norms of $F_k(\cdot, z)$, $G_k(\cdot, z)$, $f_k(\cdot, z)$ and $g_k(\cdot, z)$. Note that M , γ and w are as in Assumption 2.1(ii), and M_γ is as in (7).

Lemma C.1. *For any threshold z and integer $k \geq 0$,*

- (a) $\|F_k(\cdot, z) - F(\cdot, z)\|_w \leq M_\gamma \gamma^k$ and $\|G_k(\cdot, z) - G(\cdot, z)\|_w \leq M_\gamma \gamma^k$;
- (b) $\|F_k(\cdot, z)\|_w \leq M_\gamma$ and $\|G_k(\cdot, z)\|_w \leq M_\gamma$;
- (c) $\|f_k(\cdot, z) - f(\cdot, z)\|_w \leq 2M_\gamma \gamma^k$ and $\|g_k(\cdot, z) - g(\cdot, z)\|_w \leq 2M_\gamma \gamma^k$;
- (d) $\|f_k(\cdot, z)\|_w \leq 2M_\gamma$ and $\|g_k(\cdot, z)\|_w \leq 2M_\gamma$;

Proof. Proof. We only consider reward metrics, as the results for resource metrics follow similarly.

(a) The recursion (86) is a contraction map with modulus $\gamma < 1$ on $\mathbb{B}_w(\mathbf{X})$. See Hernández-Lerma and Lasserre [19, Remark 8.3.10]. Hence (cf. Remark 2.1(iv)) Banach's fixed point theorem gives $\|F_k(\cdot, z) - F(\cdot, z)\|_w \leq M_\gamma \gamma^k$.

(b) We first show by induction on k that $|F_k|(x, z) \leq C_k w(x)$, where $C_0 \triangleq 0$ and $C_k \triangleq M(1 + \dots + \gamma^{k-1})$ for $k \geq 1$, so $C_{k+1} = M + \gamma C_k$. For $k = 0$, such a result trivially holds. Assume now $|F_k|(x, z) \leq C_k w(x)$ for some $k \geq 0$. Then, from Assumption 2.1 and (86) we obtain

$$\begin{aligned} |F_{k+1}|(x, z) &\leq |r|(x, 1_{\{x > z\}}) + \beta \int |F_k|(y, z) Q^1(x, dy) \leq Mw(x) + \beta \int C_k w(y) Q^1(x, dy) \\ &\leq Mw(x) + \gamma C_k w(x) = C_{k+1} w(x), \end{aligned}$$

which proves the induction step. Since $C_k < M_\gamma$, it follows that $|F_k|(x, z) \leq M_\gamma w(x)$.

(c) Using (19), (88), part (a) and Assumption 2.1(ii.b), we obtain

$$\begin{aligned} |f_k(x, z) - f(x, z)| &\leq \beta \int |F_{k-1}(y, z) - F(y, z)| Q^1(x, dy) + \beta \int |F_{k-1}(y, z) - F(y, z)| Q^0(x, dy) \\ &\leq \beta M_\gamma \gamma^{k-1} \int w(y) Q^1(x, dy) + \beta M_\gamma \gamma^{k-1} \int w(y) Q^0(x, dy) \leq 2M_\gamma \gamma^k w(x). \end{aligned}$$

(d) This part follows along the same lines as part (b). □

Letting $\underline{g}(x, z) \triangleq \inf_k g_k(x, z)$, consider the following version of condition (PCLI1):

$$\underline{g}(x, z) > 0 \quad \text{for every state } x \text{ and threshold } z. \tag{90}$$

Note that, as $g_k(x, z)$ converges to $g(x, z)$ (see Lemma C.1(c)), the condition (90) implies (PCLI1).

The following result shows that the finite-horizon MP metric $m_k(x, z)$ and MP index $m_k^*(x)$ converge at least linearly with rate γ to $m(x, z)$ and $m^*(x)$, respectively.

Proposition C.1. *Let (90) hold. Then, for any state x , threshold z and integer $k \geq 0$,*

- (a) $|m_k(x, z) - m(x, z)| \leq 2M_\gamma \gamma^k w(x)(1 + |\phi|(x, z))/\underline{g}(x, z)$;
- (b) $|m_k^*(x) - m^*(x)| \leq 2M_\gamma \gamma^k w(x)(1 + |m^*(x)|)/\underline{g}(x, x)$.

Proof. Proof. (a) From

$$m_k(x, z) - m(x, z) = \frac{f_k(x, z) - f(x, z)}{g_k(x, z)} - \frac{g_k(x, z) - g(x, z)}{g_k(x, z)} m(x, z),$$

Lemma C.1(c) and $g_k(x, z) \geq \underline{g}(x, z) > 0$, we obtain the stated inequality.

(b) This part follows by setting $z = x$ in part (a). □

References

- [1] E. J. Anderson and P. Nash. *Linear Programming in Infinite-dimensional Spaces: Theory and Applications*. Wiley, Chichester, UK, 1987.
- [2] R. Bellman. A problem in the sequential design of experiments. *Sankhyā*, 16:221–229, 1956.
- [3] D. Bergemann and J. Välimäki. Bandit problems. Discussion Paper 1551, Cowles Foundation, Yale University, New Haven, CT, 2006.
- [4] D. Bertsimas and J. Niño-Mora. Conservation laws, extended polymatroids and multiarmed bandit problems; a polyhedral approach to indexable systems. *Math. Oper. Res.*, 21:257–306, 1996.
- [5] R. N. Bradt, S. M. Johnson, and S. Karlin. On sequential designs for maximizing the sum of n observations. *Ann. Math. Statist.*, 27:1060–1074, 1956.
- [6] F. Caro and J. Gallien. Dynamic assortment with demand learning for seasonal consumer goods. *Management Sci.*, 53:276–292, 2007.
- [7] F. Caro and O. S. Yoo. Indexability of bandit problems with response delays. *Probab. Eng. Inf. Sci.*, 24:349–374, 2010.
- [8] M. Carter and B. van Brunt. *The Lebesgue–Stieltjes Integral: A Practical Introduction*. Springer, New York, 2000.
- [9] E. G. Coffman, Jr. and I. Mitrani. A characterization of waiting time performance realizable by single-server queues. *Oper. Res.*, 28:810–821, 1980.
- [10] J. L. Doob. *Measure Theory*. Springer, New York, 1994.
- [11] F. Dusonchet and M. O. Hongler. Continuous-time restless bandit and dynamic scheduling for make-to-stock production. *IEEE Trans. Robotics and Automat.*, 19:977–990, 2003.
- [12] J. C. Gittins. Bandit processes and dynamic allocation indices. *J. Roy. Statist. Soc. Ser. B*, 41: 148–177, 1979. with discussion.
- [13] J. C. Gittins. *Multi-armed Bandit Allocation Indices*. Wiley, Chichester, UK, 1989.
- [14] J. C. Gittins and D. M. Jones. A dynamic allocation index for the sequential design of experiments. In J. Gani, K. Sarkadi, and I. Vincze, editors, *Progress in Statistics (Europ. Meeting of Statisticians, Budapest, 1972)*, pages 241–266. North-Holland, Amsterdam, The Netherlands, 1974.

- [15] W. R. Heilmann. Linear programming of dynamic programs with unbounded rewards. *Methods Oper. Res.*, 24:94–105, 1977.
- [16] W. R. Heilmann. Solving a general discounted dynamic program by linear programming. *Probab. Theory Rel. Fields*, 48:339–346, 1979.
- [17] O. Hernández-Lerma and D. Hernández-Hernández. Discounted cost Markov decision processes on Borel spaces: the linear programming formulation. *J. Math. Anal. Appl.*, 183:335–351, 1994.
- [18] O. Hernández-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, New York, 1996.
- [19] O. Hernández-Lerma and J. B. Lasserre. *Further Topics on Discrete-Time Markov Control Processes*. Springer, New York, 1999.
- [20] D. P. Heyman and M. J. Sobel. *Stochastic Models in Operations Research. Vol. II: Stochastic Optimization*. McGraw-Hill, New York, 1984.
- [21] B. A. Huberman and F. Wu. The economics of attention: Maximizing user value in information-rich environments. *Adv. Complex Syst.*, 11:487–496, 2008.
- [22] V. Krishnamurthy and R. J. Evans. Hidden Markov model multiarm bandits: A methodology for beam scheduling in multitarget tracking. *IEEE Trans. Signal Process.*, 49:2893–2908, 2001.
- [23] V. Krishnamurthy and B. Wahlberg. Partially observed Markov decision process multiarmed bandits — Structural results. *Math. Oper. Res.*, 34:287–302, 2009.
- [24] U. D. Kumar and H. Saranga. Optimal selection of obsolescence mitigation strategies using a restless bandit model. *European J. Oper. Res.*, 200:170–180, 2010.
- [25] B. F. La Scala and B. Moran. Optimal target tracking with restless bandits. *Digit. Signal Process.*, 16:479–487, 2006.
- [26] J. Le Ny, M. Dahleh, and E. Feron. Multi-UAV dynamic routing with partial observations using restless bandit allocation indices. In *Proc. 2008 American Control Conf.*, pages 4220–4225, New York, 2008. IEEE.
- [27] J. Le Ny, E. Feron, and M. Dahleh. Scheduling continuous-time Kalman filters. *IEEE Trans. Autom. Control*, 56:1381–1394, 2011.
- [28] S. A. Lippman. On dynamic programming with unbounded rewards. *Management Sci.*, 21:1225–1233, 1975.
- [29] K. Liu and Q. Zhao. A restless bandit formulation of opportunistic access: Indexability and index policy. In *Proc. 5th IEEE Conf. Sensor, Mesh and Ad Hoc Commun. Networks (SECON '08)*, pages 1–5, Piscataway, NJ, 2008. IEEE.
- [30] K. Liu and Q. Zhao. Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access. *IEEE Trans. Inf. Theory*, 56:5547–5567, 2010.
- [31] K. Liu, Q. Zhao, and B. Krishnamachari. Dynamic multichannel access with imperfect channel state detection. *IEEE Trans. Signal Process.*, 58:2795–2808, 2010.

- [32] K. Liu, R. Weber, and Q. Zhao. Indexability and Whittle index for restless bandit problems involving reset processes. In *Proc. 50th IEEE Conf. Decision and Control (CDC) and European Control Conf. (ECC)*, pages 7690–7696, Piscataway, NJ, 2011. IEEE.
- [33] A. Mahajan and D. Teneketzis. Multi-armed bandit problems. In A. O. Hero III, D. A. Castañón, D. Cochran, and K. Kastella, editors, *Foundations and Applications of Sensor Management*, chapter 6, pages 121–151. Springer, New York, 2008.
- [34] J. Niño-Mora. Restless bandits, partial conservation laws and indexability. *Adv. Appl. Probab.*, 33:76–98, 2001.
- [35] J. Niño-Mora. Dynamic allocation indices for restless projects and queueing admission control: A polyhedral approach. *Math. Program.*, 93:361–413, 2002.
- [36] J. Niño-Mora. Restless bandit marginal productivity indices, diminishing returns and optimal control of make-to-order/make-to-stock M/G/1 queues. *Math. Oper. Res.*, 31:50–84, 2006.
- [37] J. Niño-Mora. Dynamic priority allocation via restless bandit marginal productivity indices. *TOP*, 15:161–198, 2007.
- [38] J. Niño-Mora. An index policy for dynamic fading-channel allocation to heterogeneous mobile users with partial observations. In *Proc. NGI 2008, 4th Euro-NGI Conf. Next Generation Internet Networks*, pages 231–238, New York, 2008. IEEE.
- [39] J. Niño-Mora. A restless bandit marginal productivity index for opportunistic spectrum access with sensing errors. In *Proc. NETCOOP 2009, 3rd EuroNG Conf. Network Control and Optim.*, Lect. Notes Comput. Sci., pages 60–74, Berlin, 2009. Springer.
- [40] J. Niño-Mora. Conservation laws and related applications. In *Wiley Encyclopedia Oper. Res. Management Sci.* Wiley, New York, 2011. DOI: 10.1002/9780470400531.eorms0186.
- [41] J. Niño-Mora and S. S. Villar. Multitarget tracking via restless bandit marginal productivity indices and Kalman filter in discrete time. In *Proc. 2009 CDC/CCC, Joint 48th IEEE Conf. Decision and Control and 28th Chinese Control Conf.*, pages 2905–2910, New York, 2009. IEEE.
- [42] J. Niño-Mora and S. S. Villar. Sensor scheduling for hunting elusive hiding targets via Whittle’s restless bandit index policy. In *Proc. NETGCOOP 2011, 5th Int. Conf. Network Games, Control and Optim.*, pages 124–131. IEEE, 2011.
- [43] C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of optimal queueing network control. *Math. Oper. Res.*, 24:293–305, 1999.
- [44] M. Raissi-Dehkordi and J. S. Baras. Broadcast scheduling in information delivery systems. In *Proc. Global Telecomm. Conf. (GLOBECOM ’02)*, pages 2935–2939. IEEE, 2002.
- [45] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.
- [46] J. G. Shanthikumar and D. D. Yao. Multiclass queueing systems: Polymatroidal structure and optimal scheduling control. *Oper. Res.*, 40:S293–S299, 1992.
- [47] T. Temple and E. Frazzoli. Whittle-indexability of the cow path problem. In *Proc. 2010 American Control Conf.*, pages 4152–4158. IEEE, 2010.

- [48] J. A. E. E. Van Nunen and J. Wessels. A note on dynamic programming with unbounded rewards. *Management Sci.*, 24:576–580, 1978.
- [49] M. H. Veatch and L. M. Wein. Scheduling a multiclass make-to-stock queue: Index policies and hedging points. *Oper. Res.*, 44:634–647, 1996.
- [50] K. Wang and L. Chen. On optimality of myopic policy for restless multi-armed bandit problem: An axiomatic approach. *IEEE Trans. Signal Proc.*, 60:300–309, 2012.
- [51] R. B. Washburn. Application of multi-armed bandits to sensor management. In A. O. Hero III, D. Castañón, D. Cochran, and K. Kastella, editors, *Foundations and Applications of Sensor Management*, chapter 7, pages 153–175. Springer, New York, 2008.
- [52] R. Weber. Multiarmed bandits and Gittins index. In *Wiley Encyclopedia Oper. Res. Management Sci.* Wiley, New York, 2011. DOI: 10.1002/9780470400531.eorms1032.
- [53] R. R. Weber and G. Weiss. On an index policy for restless bandits. *J. Appl. Probab.*, 27:637–648, 1990.
- [54] Y. Wei, F. R. Yu, and M. Song. Distributed optimal relay selection in wireless cooperative networks with finite-state Markov channels. *IEEE Trans. Vehic. Tech.*, 59:2149–2158, 2010.
- [55] J. Wessels. Markov programming by successive approximations with respect to weighted supremum norms. *J. Math. Anal. Appl.*, 58:326–335, 1977.
- [56] P. Whittle. Restless bandits: Activity allocation in a changing world. In J. Gani, editor, *A Celebration of Applied Probability*, volume 25A of *J. Appl. Probab.*, pages 287–298. Appl. Probab. Trust, Sheffield, UK, 1988.
- [57] P. Whittle. *Optimal Control: Basics and Beyond*. Wiley, Chichester, UK, 1996.